D uring the past two decades, Iran has witnessed a positive trend in both input and output of medical research.\textsuperscript{1-4} The amount of expenditure in scientific research has increased from 0.55\% to 0.87\% of the GDP between 2001 and 2009, and the set target is 2.5\% to be achieved by 2015 and 4\% in 2030.\textsuperscript{5} Iran also plans to increase its education spending from 5.49\% in 2007 to over 7\% in 2030.\textsuperscript{6} The number of publications from Iran has increased from 5034 in 1996 to 20244 in 2008. This 18-fold increase outstrips that of any other country in the region.\textsuperscript{7,8} This increase in output has been accompanied by an increase in citations over the same time period.\textsuperscript{1,4} and an increase in the number of Iranian journals indexed in internationally recognized databases.\textsuperscript{9} One of the major activities at the Deputy for Research at Ministry of Health and Medical Education(MOHME) was the establishment of a regular research system, evaluation and monitoring, and an emphasis on research translation into innovation.

Research system evaluation and monitoring has an important role in the recent performance of Iranian medical universities and the Health System Research Institute. Evaluation can hold researchers and research institutes accountable for the research projects they design and launch, in addition to the research results they report. In the same context, the feedback attained through research evaluation can be applied in steering decision-making at the level of individuals, institutes, and the entire health system. This feedback can further show the strengths and weaknesses of the research system, which can justify the monetary and human resources that are spent on research. In short, it can help develop and strengthen the research infrastructure at all levels of the health system.\textsuperscript{10}

A paper by Peykari et al. presents the methodology and results of the five-year evaluation that has been, for the first time, devised and implemented in Iran. This evaluation method encompasses indicators advocated by the WHO and UNESCO and has been specifically tailored for the Health Research System (HRS) in Iran. As properly mentioned in the paper, the prominent strengths of this method can be summarized into its participatory process and its dynamism. The results of the evaluation reveal improvement in almost all indicators that have been measured in this study, from 2003 to 2008.

Evaluation methods, if not carefully designed and implemented, may lead to substantial bias in the results they release. The current method is likewise not spared from bias. The sources of bias embedded in this study originate from: 1) Methods of data collection and 2) the specific indicators that have been selected to evaluate HRS in Iran.

In the first category, the main defect is absence of the peer review process in the evaluation of HRS. Professional peer review is widely accepted and adopted in health research.\textsuperscript{11} It is the mainstay for approving research projects before their conduct and validating research results before their publication and dissemination.\textsuperscript{12} Peer review is likewise an indispensable component of research process evaluation. It is difficult for researchers, whether individually or in a team, to spot every mistake or flaw in their works. An opportunity for improvement may be more obvious to someone with special expertise. Showing research to others who look at the work with a fresh eye increases the probability that weaknesses be identified and improved. In the evaluation method described by Peykari et al., the absence of peer reviewers to independently evaluate research centers is evident.

Even more appalling is the absence of well-qualified external experts in the HRS evaluation reported by Peykari et al. There are numerous experts in this field in our country who could be independent and unbiased referees in the evaluation process. External experts are not restricted to Iranians. We could benefit from international professionals. As Iran is mounting the first steps in expanding its health research infrastructure, taking advantage of experiences that have been successful in countries with more developed HRS could considerably reduce the risk of our mistakes.

The flaw is not limited to neglecting external individual experts and peer reviewers in this study. The baseline data for HRS evaluation have been collected by medical science universities per se rather than by individual independent evaluators, and have been subsequently reported to the Deputy of Research and Technology in the Ministry of Health. It is worth noting that almost all of these evaluators are directors of their affiliated medical science universities and have been biased towards the performance of their own research body.

In the second category of sources of bias, the specific indicators selected for HRS evaluation should be discussed. In the stewardship axis, three indicators have been measured: priority setting, strategic planning, and establishment of an ethics committee. It has been announced by the Deputy of Research and Technology that medical universities will be evaluated by these three criteria. It is obvious that none of the universities would have neglected this rule and all have reported that they meet the criteria: priorities have been set, strategic plans developed, and ethics committees established. However, there is no way to qualitatively assess the actual function of a university in these domains.
Other indicators in the other two axes, capacity building and knowledge production, are similarly quantitative rather than qualitative. Capacity building indicators focus on number of workshops and congresses and do not reflect the real “capacity”, or the real environment, which may have been built for research and researchers. Indicators in knowledge production are likewise bibliometric and do not reflect the actual quality of the research products.

In conclusion, the possibility for removing these flaws should be reviewed. The first category of biases mentioned above completely relate to the processes that have been adopted for data collection. All of these defects can be eliminated through careful and precise planning. Involving peer reviewers, external independent experts, and unbiased officials in HRS evaluation will not only certify the validity of our evaluation process, it will also enrich our medical research culture and will bring us closer to the ultimate goal of our HRS evaluation, which is building a health research infrastructure. The second category of biases, the fact that indicators are mainly quantitative and evaluate knowledge production rather than translation, is not entirely the defect of this evaluation system. This defect also originates in the underdeveloped health infrastructure in our country. There is no valid way to measure what fraction of our research results are translated into improved healthcare services. We cannot ensure that research results produced in developed countries be reflected in our medical practice. Therefore, although it is necessary to guide strategies based on local research results, we may still benefit if we focus on more effective knowledge translation in our country.

Despite all of the aforementioned arguments, this first attempt aimed at evaluating the HRS in Iran can be invaluable for the enrichment of our healthcare and the health research infrastructure through its subsequent improvements.

References