Leadership in Health Services

Exploring valid and reliable assessment methods for care management education: A Delphi study

Lokke Gennissen Lorette Stammen Jolien Bueno-de-Mesquita Sietse Wieringa Jamiu Busari

Received 27 September 2015
Revised 1 April 2016
Accepted 27 April 2016

Exploring valid and reliable assessment methods for care management education

A Delphi study

Lokke Gennissen
Faculty of Medicine, Erasmus University, Rotterdam, The Netherlands
Lorette Stammen
Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, University of Maastricht, Maastricht, The Netherlands
Jolien Bueno-de-Mesquita
Platform for Medical Leadership, Den Haag, The Netherlands
Sietse Wieringa
Platform for Medical Leadership, Utrecht, The Netherlands, and
Jamiu Busari
Pediatric Residency Program, Department of Pediatrics, Zuyderland Medical Center, Heerlen, The Netherlands and Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, University of Maastricht, Maastricht, The Netherlands

Abstract

Purpose – It is assumed that the use of valid and reliable assessment methods can facilitate the development of medical residents' management and leadership competencies. To justify this assertion, the perceptions of an expert panel of health care leaders were explored on assessment methods used for evaluating care management (CM) development in Dutch residency programs. This paper aims to investigate how assessors and trainees value these methods and examine for any inherent benefits or shortcomings when they are applied in practice.

Design/methodology/approach – A Delphi survey was conducted among members of the platform for medical leadership in The Netherlands. This panel of experts was made up of clinical educators, practitioners and residents interested in CM education.

Findings – Of the respondents, 40 (55.6 per cent) and 31 (43 per cent) participated in the first and second rounds of the Delphi survey, respectively. The respondents agreed that assessment methods currently being used to measure residents' CM competencies were weak, though feasible for use in many residency programs. Multi-source feedback (MSF, 92.1 per cent), portfolio/e-portfolio (86.8 per cent) and knowledge testing (76.3 per cent) were identified as the most commonly known assessment methods with familiarity rates exceeding 75 per cent.

Leadership in Health Services
Vol. 29 No. 3, 2016
pp. 240-250

1751-1879
DOI 10.1108/LHS-09-2015-0029

The authors thank the members of PML for providing their expert opinions in this Delphi survey. The authors thank Dr Rob Moonen, Zuyderland Medical Center, for his contribution to this study.

Practical implications – The findings suggested that an "assessment framework" comprising MSF, portfolios, individual process improvement projects or self-reflections and observations in clinical practice should be used to measure CM competencies in residents.

Originality/value – This study reaffirms the need for objective methods to assess CM skills in post-graduate medical education, as there was not a single assessment method that stood out as the best instrument.

Keywords Health care, Competency, Medical education, Physician, Postgraduate, Care management

Paper type Research paper

Introduction

In recent years, health care delivery has witnessed many changes due to the rising demands for highly efficient health care systems, cost-effective health care programs and increased professional accountability from practicing physicians. The expectations of many health care funders, governments and policymakers have also increased; in addition to their primary medical knowledge, physicians should be highly skilled in other professional domains as well (Bohmer and Knoop, 2007). Many of the expected competencies now constitute the content of different educational frameworks, including the Canadian Medical Educational Directives for Specialists (CanMEDS) and the Accreditation Council for Graduate Medical Education (ACGME). Also, these expectations have been influential in the development of various national residency programs worldwide. Seven distinct roles that physicians should master upon completion of their professional training have been identified in the CanMEDS framework: collaborator, health advocate, medical expert, communicator, professional, scholar and manager/leader. The ACGME in the USA, however, describes a separate competency in its framework called systems-based practice. This skill refers to the physician's ability to participate and lead in integrated health care systems (or teams) and involves collaboration with allied health care professionals in the delivery of care. Furthermore, the management competency in the CanMEDS framework and the systems-based practice in the ACGME framework demonstrate resemblance in their focus on the organization of sustainable practices that contribute to the effectiveness of health care systems (Royal College of Physicians and Surgeons of Canada, 2005; Swing, 2007).

Although the competency of system-based practice is directed towards the efficient use of health care resources (Swing, 2007), the skill of being a manager lies in raising the individual physicians' awareness of health care systems and how they can act responsibly within these contexts (Royal College of Physicians and Surgeons of Canada, 2005). However, despite the differences and shared similarities in both cases, physicians are expected to be responsible for the organization and delivery of health care within a broader context of health care systems. Taking the observed overlap in the objectives of both the CanMEDS and ACGME competency frameworks into account, we chose the term "care management" (CM) to represent management and leadership capabilities in this paper. Hence, CM (which had been used previously in the literature) constituted our overarching construct in describing the combined competencies of manager and system-based practice (Halpern *et al.*, 2001; Busari *et al.*, 2014).

In the past couple of years, the development of CM competencies in (trainee) physicians has received much attention. Consequently, several post-graduate medical

institutions have engaged more in investigating the impact of CM training programs within their curricula (Bax *et al.*, 2011; Brouns *et al.*, 2010; Daugird and Spencer, 1990; Patel *et al.*, 2009). So far, the available data reflecting the impact of these initiatives have been gathered indirectly. Some of the sources of these data include information from the number of trainees attending workshops, residents' ratings of training programs and, in some cases, from pre- and post-test assessments of trainees' knowledge in CM (Busari *et al.*, 2014). Unfortunately, it has been difficult to draw conclusions from these sources, as an increase in perceived knowledge does not necessarily reflect an improvement in the behavior and management capabilities of physicians. Therefore, the question that one should ask is "to what degree is a trainee able to apply the acquired set of CM competencies in practice after participating in any of these programs?".

In a recent study by Stutsky *et al.* (2012), it was shown that physicians considered the seven CanMEDS roles to be at least moderately important for their practice(s) (Stutsky *et al.*, 2012). Recent studies also revealed that the domain of management and leadership development lacked the attention it deserved within the post-graduate medical curricula (Bax *et al.*, 2011; Brouns *et al.*, 2010; Daugird and Spencer, 1990; Patel *et al.*, 2009). Although the need for CM competency development among physicians is essential and recognized as such (Busari *et al.*, 2011), many residency programs are still struggling to incorporate CM education formally into their curricula as well as searching for better strategies to assess CM development in trainees.

The literature on assessment in (medical) education suggests that assessment, in general, tends to influence the learning process of trainees (Epstein, 2007). Hence, it is reasonable to assume that when assessment methods are valid and reliable, they will contribute to the quality of learning in training programs, including CM training. Within this line of thinking, reliable and valid assessment methods that target specific domains of management and leadership would be necessary to evaluate the efficacy of any CM training program. Furthermore, the utilization of these methods could be a potential strategy for designing effective management training programs.

Gap analysis

A recent 2012 study that we conducted identified the need for widespread reliable and valid assessment methods in CM (Busari *et al.*, 2014). The review showed that evidence for the validity and reliability of the assessment methods was lacking and that there was a gap in the knowledge of those methods currently being used to assess CM in residents. As a result, we decided to perform a qualitative exploration of current assessment methods for CM in post-graduate training settings. We were particularly interested in identifying specific assessment methods that were being used in the clinical workplace and highlighting any perceived benefits and shortcomings of these methods. We were also interested in detecting the essential elements that (could) constitute a valid and reliable assessment method in this domain. Our long-term objective for embarking on this exercise was to develop an assessment method that could reliably determine trainee physicians' knowledge and skills in CM.

Methods

Study design

We chose to use the Delphi method for this study, as we intended to obtain and synthesize the opinions of respondents whom we considered experts in the field. The

Delphi method is a structured, interactive and systematic investigative approach that relies on the answers or forecasts of a panel of experts to a set of predefined questions. The method assumes that estimates (or decisions) from a structured group of individuals are more accurate than those from unstructured groups. Conducted in at least two rounds, an analysis of the responses from the panel is performed and summarized after each round. By subjecting the set of questions to a couple of reiterations, any variance from the answers within the group is expected to diminish and converge towards a "correct" answer during the process. The anonymous summaries from a previous round together with new supporting arguments are presented to each expert again in the next round of review. However, when a pre-defined stop criterion is met, for example, "*x*" number of rounds, achievement of consensus or stability of results the process is discontinued. Subsequently, the mean or median scores of the final round are used to determine the outcome (Rowe and Wright, 2001).

Participants

The respondents who participated in our survey were members of the (Platform Medisch Leiderschap, PML) in The Netherlands. The PML was created in 2012 and is a Dutch collaboration of health care professionals with a particular interest in medical management and leadership. At the time of the survey, the PML consisted of 72 members, 13 medical residents, 14 health care management professionals and 45 medical specialists and or clinical educators. We chose to conduct this survey with this group of respondents because we considered them to be ideal and well suited to provide us with the information we needed on CM development. This was supported by their involvement in training programs (as both trainees and trainers) and their knowledge of current methods of assessment in CM training. Prior to conducting the survey, ethical approval was sought and granted by the Ethical Review Board of the Zuyderland Medical Hospital, as the study did not constitute any risk to the privacy of respondents or patients.

Procedure

In October 2012, all the members of the PML were invited to participate in the Delphi survey. The participants who accepted the invitation gave informed consent prior to the study. All members received an invitation for the survey by e-mail, followed by a reminder after two weeks. In total, 40 members completed the first version of the Delphi survey, with a response rate of 55.6 per cent. The response rate in the second round was lower (43 per cent) and included 31 respondents.

We defined the following two research themes for this survey:

(1) the identification of assessment methods for current use; and

(2) satisfaction with the currently used assessment methods based on clinical experience and evidence from the literature (Busari *et al.*, 2014).

We were aware that many assessment methods had been developed to evaluate management competencies in post-graduate medical training. Therefore, the first theme was designed to explore which assessment methods our expert panel members were aware of and to identify the methods currently in use in daily clinical practice. The second theme focused on the satisfaction of educators and trainees regarding the use of current assessment methods for summative and formative purposes. The questions for this theme included those investigating the benefits and

shortcomings of assessment methods, the ability to make decisions about the competence level of residents and the value of evaluation methods for giving constructive feedback to residents. We were particularly interested in knowing whether clinicians or educators experienced difficulties in measuring trainees' management competencies and if there was a need for valid assessment methods to measure this competence.

Data collection

We conducted a two-round Delphi survey with the selected group of respondents whom we invited for both rounds of the survey by e-mail. The email invitations also contained a link to an electronic questionnaire. The first Delphi round consisted of a 17-item questionnaire in which participants could express their opinions about specific questions on a four-point Likert scale (agree, somewhat agree, somewhat disagree and disagree). An example of such a question is "the assessment method most frequently used by me gives the resident(s) a clear view of the aspects of CM that require improvement/progress/development". The participants were also asked to indicate their preferred instruments of choice from a list of assessment methods in the questionnaire. Examples of the methods included oral examinations, simulated patient contacts, objective structured clinical examination multi-source feedback (MSF) and portfolios. There was also the opportunity for the respondents to suggest alternative assessment methods if they felt that a particular method was missing from the list. They also had the chance to provide written comments should they wish to clarify or supplement any of their responses. The second Delphi survey was conducted one month after the first round and contained eight major questions. We designed the items in the second Delphi round in such a way that the respondents could respond with a simple "yes or no" or "right/wrong" answer. An example of a question that we used in the second Delphi round is "the assessment methods currently available do not measure CM competencies adequately".

Data analysis

In our survey, "consensus" was defined as being achieved when a clear majority of our respondents agreed with a statement, theme or research question(s). We determined *a priori* clinically relevant consensus as more than two-thirds (>67 per cent) of the participants agreeing or disagreeing with a theme or research question. Consensus, therefore, implied that a statement was either accepted or declined by the (majority of the) group. Reports that did not reach consensus were edited and rephrased so that they were aligned with the comments provided by the respondents for the next Delphi round. Comments that did not reach consensus were those in which the respondents' opinions were diverse or achieved less than 67 per cent agreement. The point at which saturation was reached was defined as the criterion for ending the Delphi round.

Results

Characteristics

The demographic characteristics of the PML members were diverse; 21 female and 19 male participants completed the Delphi survey. The participants' ages ranged from 24-60 years, with a mean of 41 years. Additionally, the participants' job descriptions represented different work domains such as primary care facilities (32.5 per cent); academic- (25 per cent) and community-based hospitals (15 per cent); universities (5 per cent); and other health care services (22.5 per cent). Finally, majority of the respondents

had more than six years of experience in health care (6-10 years = 23.7 per cent; >10 years = 42.1 per cent).

First round. The first set of questions in the Delphi survey focused on the participants' familiarity with methods used for assessing residents' management competencies. We provided the respondents with an overview of commonly used assessment methods in the area of residency training such as MSF, portfolios, knowledge tests, clinical observations, self-assessments and simulated patient consultations. MSF (92.1 per cent), portfolio/e-portfolio (86.8 per cent) and knowledge testing (76.3 per cent) were the most commonly known assessment methods with familiarity rates exceeding 75 per cent. Three respondents made use of the explanatory notes to supplement our list of assessment methods with oral examinations, assessment of implementation plans, coaching on the job and a combined module with practical assignments.

One of the participants also mentioned progress interviews and reflection reports based on a frequently used individual educational plan for trainees in The Netherlands (Koninklijke Nederlandse Maatschappij tot bevordering der Geneeskunst). Also, the National Health Service (NHS) framework was referred to in multiple explanatory notes. Half of our respondents stated that the topic of CM received attention in the workplace setting. Interestingly, only 10 per cent of the respondents in the resident sub-group agreed with the statement that CM received attention in the workplace setting, in contrast to 64.3 per cent in the group containing physicians and managers. The statements regarding CM education resulted in more homogenous data, with a majority of 64.5 per cent stating that there was insufficient attention given to CM education.

Validity/reliability/feasibility

The results from the first round of our Delphi study were suggestive of a lack of validity in the methods being used to assess CM competencies. Statements about their usefulness in (current) practice showed that the respondents agreed on the feasibility of the assessment methods. The findings also revealed that the methods in current use were perceived as not being reproducible and incapable of assessing all aspects of CM competencies. Furthermore, they are unable to provide a good reflection of residents' management capabilities.

Second round. The participants' responses in the second Delphi round reaffirmed that current assessment methods were incapable of providing clear insights into trainee performance and were not reproducible. Although they felt that the methods were unable to differentiate between the different competency levels of residents, they agreed that they did provide useful feedback. Two assessment methods that were acknowledged as valid instruments were the MSF and portfolio. Other methods that were also considered feasible for assessing management competencies included the short clinical examinations, computerized simulations, simulated patient consultations, portfolios and self-assessments (reflections).

Explanatory notes

Analysis of the content of the explanatory notes revealed that a few participants expressed concern about the definition of CM despite the fact that we provided them with an introduction and definition of CM in their invitation letters. Their main critic was the lack of an explicit description of CM and its components. The explanatory notes also revealed that there was a need to explore the development of additional methods

and investigate the current methods of assessment. For example, one participant said: " all of the suggestions are worth investigating […]. for the development of (reliable) assessment methods".

A recurring theme that emerged from the analysis of the explanatory notes was the perception that management competencies should be measured using a combination of assessment methods. For example, one participant made a remark about the importance of using triangulation to obtain more reliable measurements, whereas another mentioned that a key feature of an assessment method should be its ability to provide an entrée for a feedback conversation. Triangulation can be described as the combination of different strategies within a study to answer a particular research question (Thurmond, 2001). When this process involves the combination of two or more "data sources", however, it is referred to as data source triangulation. Triangulation could also involve the combination of different investigators, theoretical perspectives, methodologic approaches or analytical methods. In such cases, it is usually typified after the source, for example, investigator triangulation or theoretical investigation (Kimchi *et al.*, 1991). Hence, going back to the context of the respondent's recommendation, it would be logical to assume that data source or methodological triangulation would be the potential strategies to obtain reliable measurements (Thurmond, 2001).

There was also a suggestion to delegate the assessment of the management competencies to managers. The overall recommendation of the expert panel was to design an assessment framework made up of MSF, portfolio and direct clinical observation by supervisors for assessing CM competency in specific clinical situations.

An overview of topics addressed by the Delphi surveys is provided in Table I.

Discussion

The objective of this study was to explore how CM was being assessed in current practice and to identify perceived advantages and shortcomings of the assessment methods. We were interested in discovering if we could reliably determine the managerial competencies of trainee physicians and provide them with individual insights to improve specific areas that required improvement whilst monitoring the progress of their development in this domain. We made use of a Delphi approach to investigate if assessment methods were available that could reliably measure CM competencies of trainee physicians in post-graduate medical education. For this, a panel of practicing health professionals and trainees with specific expertise in CM was recruited.

The results revealed that there were several assessment methods available for assessing CM competencies. However, there was not one method that stood out as the best instrument for assessment as perceived by our respondents. This outcome was in line with that from a literature review we conducted in 2012 that appraised assessment methods currently being used in CM education. In that study, none of the articles included in this review showed any evidence of reliability in the assessment methods they reported. Furthermore, it was difficult to differentiate between what was being measured in terms of "competence" or "competency" in CM (Busari *et al.*, 2014).

In addition to the aforementioned points, there were additional findings that we discovered in our Delphi study. Firstly, our expert panel agreed that there was a need for an objective method for the assessment of CM competencies. Secondly, the current assessment methods failed to discriminate areas for improvement and, hence, were of limited use as formative feedback methods for residents. This finding was a significant

Topic/theme addressed by the items in the questionnaire	Round I	Round II
General characteristics of the expert panel	Items 1-6	–
Attention for management tasks at the workplace	Item 7	Item 1
Familiarity with assessment methods	Item 8	–
Validity/Reproducibility/Feasibility of assessment methods for CM competency in general	Item 9	Item 2 Item 5
Clinical experience in using assessment methods in the past	Item 10	–
Assessment methods for CM competency currently used	Item 11	Item 3
The presence of a valid assessment method for assessment of CM competency in general	Item 12	Item 4
Presence of feasible assessment methods for assessment of CM competency	Item 13	Item 4
Perceived need for a valid assessment method for CM competency	Item 14	Item 4
The purpose of current assessment methods with respect to reproducibility, formative feedback, summative feedback and differentiation between trainees	Item 15	Item 5
Validity of assessment methods currently used for assessment of CM competency	Item 16	Item 6
Feasibility of assessment methods currently used for assessment of CM competency	Item 16	Item 7
Reproducibility of assessment methods currently used for assessment of CM competency	Item 16	Item 8
What combination of assessment methods would you recommend for assessing CM competency in (trainee) physicians	Item 17	–

Notes: This table shows the structure of the Delphi rounds; the questions that did not reach consensus (i.e. <67% agreement) were rephrased and repeated in Round II

Table I. Outline of the questionnaire items used for the Delphi survey

shortcoming when considered within the context of the assessment of learning as opposed to the evaluation of learning (van der Vleuten and Schuwirth, 2005). Thirdly, unexpected outcome from one of our sub-group analysis indicated that the opinions of residents and senior staff differed with respect to the provision of management training in the clinical workplace. The residents felt there was less attention devoted to CM development on the work floor, whereas the senior staff thought that the attention given to CM development in practice was sufficient. This discrepancy could probably mean that residents were unable to recognize CM in the way that it was being delivered in clinical practice. It could also mean that more explicit, formal and structured educational formats are required to reduce the gap between residents and physicians' perspectives.

More specifically, MSF, portfolios, knowledge tests, clinical observations, self-assessments and simulated patient consultations were the most frequently reported assessment methods used individually to measure management competencies. Of all

these methods, however, the MSF was the most commonly used. Although some the assessment methods were thought to be feasible, arguments supporting them were not strong enough to dispel the need for an instrument that was more valid, reliable and viable for assessment purposes. The NHS framework was also mentioned a couple of times by the expert panel as a reliable assessment method.

Although this framework provides an excellent overview of the different stages in which competencies were divided, there is still no known guideline that can inform us about the ideal assessment method and, thus, about how to correctly determine the stage (milestone) at which a resident is at (NHS). Therefore, in the absence of a single ideal evaluation method, the expert panel recommended the aggregation of different assessment methods into a unique assessment framework to measure CM competencies in residents, namely, MSFs, portfolios, observations in clinical practice and individual process improvement projects or self-reflection journals.

Limitations

The main advantage of this Delphi study was the collective expertise of the panel members we recruited for the survey. All the participants were familiar with assessments in medical education and worked in different fields of health care delivery or education. They were also actively involved in various domains of CM and, therefore, were capable of providing an excellent overview of the way the competency is currently being implemented and assessed in post-graduate medical education. The narrative comments that we collated provided us with rich data that enabled us to identify significant shortcomings in the currently used assessment methods. Reformulating the questions in the second round also gave us the opportunity to test whether the majority shared these assumptions. Although we only approached one group of experts in this study, we still feel that despite the small sample size, the diversity in demographic characteristics, the variation in their clinical background and the way that these were combined made the composition of the panel highly valuable. This aligns with findings in the literature that purport the importance of focussing on the quality of participants in qualitative studies, as this tends to contribute to the richness of the retrieved content and variety of information, especially when sample sizes are small (Flyvbjerg, 2006). Furthermore, as far as we know, this is the first time that an attempt is being made to provide objective and practical recommendations on how to assess residents' CM competencies in post-graduate medical education.

There are certainly a few potential caveats that we need to bear in mind. For example, although a response rate of 55.6 per cent is high for research purposes, the size was not as large as we anticipated from the expert group. This low number could be because of reactive resistance to the number of questions asked or to the number of electronic mails health care professionals process on a daily basis. We also received less narrative feedback in the second Delphi round compared to the first. This poor response could have been due to a loss of interest among the panelists in having to repeat their views a second time. That notwithstanding, the quality and the diversity of the background of the panel members were broad enough for us to discard the possibility of any serious bias in our results.

Another point about the narrative comments was the suggestion that a few questions in the second round overlapped with some in the first Delphi round. This was not surprising though, as we had to rephrase those items that did not achieve consensus

during the first round and included them again in the second round. However, despite their comments about overlapping items in the questionnaire, the data we obtained from the two rounds were valuable enough for analysis and contained enough content to generate recommendations for CM assessment methods.

Lastly, although the selection of the expert panel was based on professional experience and diversity in their domains of expertise, one needs to exercise restraint in generalizing these findings, as the participants were either practicing or training within The Netherlands. Still, we believe that our findings can be of value to a broader international community, as the framework for post-graduate medical education in The Netherlands evolved from CanMEDS. We also think that the implications and recommendations derived from this strategically chosen group of experts can be applied to other contexts beyond the Dutch health care setting.

Conclusions

The findings from our study reaffirm the need for objective methods to assess CM skills in post-graduate medical education. However, as assessors and trainees could not identify a single method to achieve this, they proposed the use of a CM assessment framework as an alternative strategy. Such a framework will need to be made up of different assessment methods and combined in way to ensure that CM competencies can be determined objectively in clinical practice. Our findings showed that the best combination of assessment methods for such a framework include MSF, portfolios, clinical observations by supervisors and either an individual project or self-reflections. Furthermore, the choice of these assessment methods was based on the opinions of the expert panel and not on the inherent proof of validity or reliability of the individual assessment method. Because of the limitations of feasibility, we would advise that the application of these methods occur linearly within specified periods during training. Finally, we suggest that these methods should be combined and aggregated over time so that they can provide residents with summative judgments about the level of their development as they progress through their training.

Ethical approval

Ethical approval was requested and obtained from the Ethical Review Board of the Zuyderland Medical Hospital.

References

Bax, M.F.P., Berkenbosch, L. and Busari, J.O. (2011), "How do medical specialists perceive their competency as physician-managers?", *International Journal of Medical Education*, Vol. 2, pp. 133-139.

Bohmer, R.M.J. and Knoop, C.I. (2007), *The Challenge Facing the US Healthcare Delivery System*, Harvard Business School, Boston, MA, pp. 606-609.

Brouns, J.W., Berkenbosch, L., Ploemen-Suijker, F.D., Heyligers, I. and Busari, J.O. (2010), "Medical residents perceptions of the need for management education in the postgraduate curriculum: a preliminary study", *International Journal of Medical Education*, Vol. 1, pp. 76-82.

Busari, J.O., Berkenbosch, L. and Brouns, J.W. (2011), "Physicians as managers of health care delivery and the implications for postgraduate medical training: a literature review", *Teaching and Learning in Medicine*, Vol. 23 No. 2, pp. 186-196.

Busari, J.O., Stammen, L.A., Gennissen, L.M. and Moonen, R.M. (2014), "Evaluating medical residents as managers of care: a critical appraisal of assessment methods", *Advances in Medical Education and Practice*, Vol. 5, pp. 27-37.

Daugird, A.J. and Spencer, D.C. (1990), "The perceived need for physician management training", *Journal of Family Practice*, Vol. 30 No. 3, pp. 348, 351-352.

Epstein, R.M. (2007), "Assessment in medical education", *New England Journal of Medicine*, Vol. 356, pp. 387-396.

Flyvbjerg, B. (2006), "Five misunderstandings about case-study research", *Qualitative Inquiry*, Vol. 12 No. 2, pp. 219-245.

Halpern, R., Lee, M.Y., Boulter, P.R. and Phillips, R.R. (2001), "A synthesis of nine major reports on physicians' competencies for the emerging practice environment", *Academic Medicine*, Vol. 76 No. 6, pp. 606-615.

Kimchi, J., Polivka, B. and Stevenson, J.S. (1991), "Triangulation: operational definitions", *Nursing Research*, Vol. 40 No. 6, pp. 364-366.

Patel, M.S., Lypson, M.L. and Davis, M.M. (2009), "Medical student perceptions of education in health care systems", *Academic Medicine*, Vol. 84 No. 9, pp. 1301-1306.

Rowe, G. and Wright, G. (2001), "Expert Opinions in Forecasting, Role of the Delphi technique", in Armstrong, J.S. (Ed.), *Principles of Forecasting: A Handbook of Researchers and Practitioners*, Kluwer Academic Publishers, Boston, MA.

Royal College of Physicians and Surgeons of Canada (2005), *CanMEDS 2005 Framework*, available at: www.royalcollege.ca/portal/page/portal/rc/common/documents/canmeds/framework/the_7_canmeds_roles_e.pdf (accessed 21 April 2015).

Stutsky, B.J., Singer, M. and Renaud, R. (2012), "Determining the weighting and relative importance of CanMEDS roles and competencies", *BMC Research Notes*, Vol. 5, p. 354.

Swing, S.R. (2007), "The ACGME outcome project: retrospective and prospective", *Medical Teacher*, Vol. 29 No. 7, pp. 648-654.

Thurmond, V.A. (2001), "The point of triangulation", *Journal of Nursing Scholarship*, Vol. 33 No. 3, pp. 253-258.

van der Vleuten, C.P. and Schuwirth, L.W. (2005), "Assessing professional competence: from methods to programmes", *Medical Education*, Vol. 39 No. 3, pp. 309-317.

Further reading

Koninklijke Nederlandse Maatschappij tot Bevordering der Geneeskunst (2016), *Modernisering Medische Vervolgopleidingen: Individueel Opleidingsplan – Handleiding en Voorbeeld Individueel Opleidingsplan [Manual for Developing an Individual Educational Plan]*, available at: www.medischevervolgopleidingen.nl/rubrieken/opleidingsplan-en-toetsing/individueel-opleidingsplan/ (accessed 05 June 2016).

Platform Medisch Leiderschap (2016), available at: http://platformmedischleiderschap.nl/ (accessed 05 June 2016).

The National Health Service (2016), *Leadership Framework*, available at: www.leadershipacademy.nhs.uk/resources/healthcare-leadership-model/ (accessed 05 June 2016).

Wikipedia Delphi Method (2016), available at: https://en.wikipedia.org/wiki/Delphi (accessed 05 June 2016).

Corresponding author
Jamiu Busari can be contacted at: jamiu.busari@maastrichtuniversity.nl