Social Science & Medicine 220 (2019) 292–300

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Measuring management practices in India's district public health bureaucracy

Timothy Powell-Jackson[a,*], Bhaskar Purohit[b], Deepak Saxena[b], Mahaveer Golechha[b], Camilla Fabbri[a], Partha Sarthi Ganguly[b], Kara Hanson[a]

[a] Department of Global Health and Development, Faculty of Public Health and Policy, London School Hygiene and Tropical Medicine, 15-17 Tavistock Place, London, WC1H 9SH, UK

[b] Indian Institute of Public Health Gandhinagar (IIPHG), Gandhinagar-Chiloda Road, Lekawada, P.O. CRPF Camp- 382042, Gandhinagar, Gujarat, India

ARTICLE INFO

Keywords:
India
Management practices
District health system
Measurement
Reliability
Validity
Psychometric properties

ABSTRACT

Weak management is widely recognised as a key impediment to scaling-up coverage of health interventions and ensuring health systems are responsive to population needs. Yet there is scant evidence linking management practices in the public administration to effective health service delivery. We report on the development of a tool to measure management practices in India's district health bureaucracy. We first developed a conceptual framework based on a review of the literature and qualitative interviews with district public health managers. Across 16 management practices, we then drafted and piloted questions to be used with a scoring grid to evaluate process-orientated management practices. We implemented the tool in 34 districts of Maharashtra between April and July 2016, interviewing up to three district public health managers per district (n = 99). Using rigorous psychometric methods, we assessed the acceptability, reliability and validity of the tool. We present three key findings. First, the tool was feasible to implement, response rates were high, and there were no missing data. Second, internal consistency of the tool was high and test-retest reliability was comparable with other management tools used in the literature. Third, there was evidence of validity. The number of staff with a management qualification was positively associated with better management practices. Factor analysis showed that one principal component loaded positively on all the management practices although there was little support for management sub-scales. These findings provide novel evidence on the psychometric properties of a tool designed to measure management practices in the public administration of a developing country. Our framework and tool provide the basis to examine associations between district health management practices and health service delivery, and test the effectiveness of management strengthening interventions in India's public health sector.

1. Introduction

Weak public service delivery has long been regarded as a key obstacle to ensuring widespread coverage of essential health interventions (Travis et al., 2004). Nowhere is the issue more salient than India, where the state has struggled to deliver basic public services to its population, despite the presence of elite national institutions and a highly educated top brass of public administrators (Pritchett, 2009). This is perhaps best reflected in the absenteeism of frontline health workers in the public sector, estimated at 40 percent (Muralidharan et al., 2011).

One reason for the current situation is possibly poor managerial quality in the public administration. Weak management is widely recognised as a key impediment to scaling-up coverage of health interventions (Mangham and Hanson, 2010) and ensuring health systems are responsive to population needs (de Savigny and Adam, 2009). Yet there is a dearth of evidence linking practices in the public administration to effective service delivery and outcomes (Goldfinch et al., 2012). With a few exceptions (Rasul and Rogger, 2018), much of the literature on management has focused on private firms or service delivery organisations, such as hospitals, schools and universities (Bloom et al., 2014; Bloom and Van Reenen, 2007; McCormack et al., 2014).

In health systems and policy research, management has received little attention. Management is almost absent in commonly used frameworks on health system performance, falling under broader notions of “stewardship” and “governance” (WHO, 2007b). Most research on

* Corresponding author.
E-mail address: Timothy.Powell-Jackson@lshtm.ac.uk (T. Powell-Jackson).

https://doi.org/10.1016/j.socscimed.2018.10.031
Received 11 April 2018; Received in revised form 17 October 2018; Accepted 31 October 2018
Available online 02 November 2018

health systems is framed around the "hardware", with limited attention given to management and other health system "software" that shape the delivery of health services (Sheikh et al., 2011). One notable exception is recent qualitative research on the everyday resilience of district health systems and managerial responses to challenges (Gilson et al., 2017).

We embarked on a project that aimed to study the relationship between management practices in the district health bureaucracy and the delivery of health services in India. As part of the research, we developed a tool to quantitatively measure management practices across the district health offices of Maharashtra. In this paper, we present a detailed account of how the tool was developed and report on its reliability and validity. These methodological details are important because of the inherent challenge of assessing a multidimensional concept such as management in complex public sector organisations. We focused on the district because it is the primary unit entrusted with implementing policy and ensuring effective primary care service delivery in India.

2. Literature on management

2.1. Defining management

There are various bodies of the literature that speak to the question of how to conceptualise management. The first body of literature presents various frameworks of how to classify management functions in the health sector. A useful starting point is the definition given by Vriesendorp et al. (2010) in which health management is described as continuously developing the potential of an organisation to transform human and financial resources and other inputs into improved services and better health. The literature consistently emphasises two dimensions: managing (planning and using resources efficiently to produce intended results) and leading (mobilising others to envision and realize a better future) (Daire et al., 2014; Vriesendorp et al., 2010). Management practices are the set of processes of planning, budgeting, organising, staffing, controlling and problem solving (Dorros, 2006; Kotter, 2001), in relation to the management of governance, human resources, financial resources, supplies and medicines, and information (Vriesendorp et al., 2010). The World Health Organisation delineates four dimensions of good leadership and management (number of well-trained managers, competencies, support systems, enabling work environment), making a distinction between mangers and management practices (Egger et al., 2005; WHO, 2007a).

Of particular relevance to our study on district management is the idea of decision space, developed by Bossert (1998) in the context of decentralisation. The decision space approach provides a useful mapping of the functions and degrees of choice that might be transferred to local officials in the process of decentralisation. Functions include financing, service organisation, human resources, targeting and governance. A distinction is made between decision space that is governed by laws and regulations and actual or informal decision space that is defined by lack of enforcement such that officials can bend the rules or operate outside the limits of their authorised decision space.

A second body of empirical research is relevant because of the conceptualisation of management underpinning the measurement tools. The World Management Survey (WMS) represents the first systematic effort to collect data on representative samples of organisations and firms. The survey methodology, pioneered by Bloom and Van Reenen (2007), seeks to measure management in three broad areas – monitoring, targets, and incentives – with a focus on process-orientated practices that are universally considered "good". The conceptualisation of management in this way has been influenced by notions of lean manufacturing techniques, key performance indicators, and best practice relating to promotion decisions. Some of the practices share similarities with other tools used to evaluate human resource management practices (MSH, 2012).

The WMS methodology provided the starting point for a recent study of management practices in the Nigerian civil service (Rasul and Rogger, 2018). Adaptation of the assessment tool from private to public organisations took into account long-held perspectives on the importance of autonomy and delegation in public administration (Rose-Ackerman, 1986) as well as insights from 'new performance management', 'new public management' and 'good governance' agendas (Goldfinch et al., 2012; Mills et al., 2001; Minogue et al., 2000). The public administration literature highlights two broad dimensions: autonomy of middle and lower tier bureaucrats; and performance based incentives. Extensions of the tool to cover management practices relevant to public sector management included adding additional questions on capacity building, flexibility, roles and delegation. This idea of autonomy – capturing the extent to which civil servants input into policy development and implementation processes – is closely related to Bossert's (1998) concept of decision space.

A third set of studies evaluating management interventions in health care provide some insights on how management has been conceptualised. An early study in the Gambia sought to strengthen district management by improving skills and resource management through better planning and coordination (Conn et al., 1996). An important insight was that effectiveness of the intervention was limited by the degree to which decision making was centralised. More recent studies have been conducted in Kenya (Seims et al., 2012) and Zambia (Mutale et al., 2017). These interventions drew heavily on the guidance given in Vriesendorp et al. (2010), focusing on both leadership and management.

As a final remark, it is worth recognising that there is a broader literature on organisational performance that emphasises the importance of values and organisational culture (Gilson and Erasmus, 2004). In fact, much empirical work has been done in the measurement of organisational culture in health care, with a systematic review identifying thirteen instruments that have been used, typically in high-income settings (Scott et al., 2003).

2.2. Measuring management

The WMS tool has been used to collect data on management practices in firms worldwide (Bloom et al., 2014). The initial focus was on private manufacturing firms but the tool has subsequently been adapted to the measurement of management in hospitals, schools, universities, and the retail sector. Most recently, it has also been adapted to measure management practices of frontline service providers (Lemos and Scur, 2016) and civil service organisations in developing countries (Rasul and Rogger, 2018). Other management measurement tools have been used in various low- and middle-income countries (LMIC) (McKenzie and Woodruff, 2016; Seims et al., 2012).

Survey methods used to measure management vary in terms of the type of interview, the nature of the questions and the scoring method. The standard WMS method is to interview respondents over the phone (Bloom and Van Reenen, 2007). The benefit of this method is that it is possible to blind the interviewees to the fact that they are being assessed and to blind the interviewer to the characteristics of the organisation. In certain contexts, telephone interviews are challenging and face-to-face interviews are preferable, as has been done when the WMS tool has been applied in LMICs to the public sector (Lemos and Scur, 2016) and small firms (McKenzie and Woodruff, 2016). We are aware of one study that mailed the questionnaire to respondents for self-completion (Bloom et al., 2016a). Most studies interview an individual respondent, although group face-to-face interviews were used in the study of civil servants in Nigeria (Rasul and Rogger, 2018).

There is variation in the nature of the questions used to elicit information on management practices in organisations (Bloom et al., 2016b). The WMS methodology uses mostly open-ended questions whose answers provide the basis with which to evaluate the management practice (Bloom and Van Reenen, 2007). The open-ended

approach helps deal with the fact that some of the concepts are complex and require discussion and examples, but it does place more of a burden on the interviewer in terms of interpreting the responses. Other studies have used closed ended questions with answers recorded with categorical response options (Bloom et al., 2016a; McKenzie and Woodruff, 2016).

Converting responses from interviewees to a score is an important aspect of the methods. In the WMS methodology, the interviewer uses a scoring grid containing descriptors for each score to evaluate management practices (Bloom and Van Reenen, 2007). In the adaptation of the WMS tool to the public sector of developing countries, a similar method is used except that two interviewers score the management practices after the end of the interview based on their notes and reconcile any differences (Lemos and Scur, 2016). Studies that have used closed-ended questions must rank the responses from worst to best and then assign a value to each response. Finally, some studies have used subjective Likert scales to score responses (Bruhn et al., 2018). The standard approach to generating an overall score or index of management is to take an unweighted average of each item.

Tools used to measure management have not been well validated. The test-retest reliability of the WMS tools has been assessed in some studies, although imperfectly (Bloom and Van Reenen, 2010a). Second interviews have been conducted with different managers within the same organisation but by different interviewers. Results from telephone interviews have been compared with those from face-to-face interviews of other managers within the same organisation (Grous, 2011). As a form of validation, numerous studies have examined associations between management practices and performance of the organisation.

3. Methods

3.1. Study setting

The study was conducted in the state of Maharashtra, India's second most populous state with 112 million people, of which 12.4 million live in Mumbai, according to the Indian Census 2011 (Office of the Registrar General and Census Commissioner, 2013). The state is divided into 6 divisions and 35 districts. GDP per capita in the state is 134,081 Rs (USD $2090) compared with the country average of 86,879 Rs (USD $1354), making it one of the richer states in India. Under-five mortality is 29 deaths per 1000 live births and maternal mortality is 68 deaths per 100,000 live births (International Institute for Population Sciences and Macro International, 2016; Office of the Registrar General, 2016).

The Government of Maharashtra has adopted a decentralised structure similar to most other Indian states. Even though health is a state subject, most states follow a similar pattern of health care administration and management. This is largely because of a common planning framework, which is governed by the Planning Commission and the National Development Council, as well as the legacy of a common history of British colonial rule that laid the foundations of the health service bureaucracy. Further, the fiscal devolution of resources is determined by the central government and this is done through programmes, which are usually uniform across states.

The Ministry of Health and Family Welfare (MoHFW) in Maharashtra is divided into two departments, the Public Health Department and the Department of Medical Education and Drugs (see Appendix 1). Both these departments have a separate Minister and Minister of State and their Secretariat, as well as technical wings called Directorates. The district plays a key role in overseeing the delivery of rural health services and the implementation of health programmes.

The district health office is run by the District Health Officer (DHO) and a team comprising the Additional DHO, District Program Manager (DPM), district level programme managers and various support staff who administer the primary health care system of primary health centres and sub-centres.

The study received ethical approval from the Institutional Ethics Committee of the Indian Institute of Public Health in Gandhinagar (ref: TRC-IEC No. 31/2014) and the London School of Hygiene and Tropical Medicine, UK (ref: 8784).

3.2. Tool development

3.2.1. Conceptualising good management

Developing a tool to measure management requires having an idea of what good management looks like. Informed by the literature review, a number of principles guided our conceptualisation of management. First, we focused on management practices, rather than managers. While there is a literature on measuring leadership (Bandiera et al., 2017), we were primarily interested in organisational processes. Hence we did not attempt to capture what personality traits or qualities, such as inspirational leadership, constitute a good manager (Sharma and Tarp, 2018). This was also for the pragmatic reason that information on processes and systems should be possible to capture given that there is an established literature on measuring management. Second, we gave priority to management practices for which there was some consensus on what constitutes good and bad management. In other words, we sought to evaluate, and not simply describe, the management practices in place. There is of course scope for legitimate debate on individual practices. The evidence is mixed, for example, on whether targets are motivating or demotivating and whether they encourage gaming behaviours that are detrimental to organisational performance (Cleary et al., 2013; Hood, 2006).

We were cognisant of the need to have a thorough understanding of the district public health bureaucracy in India, particularly its organisational structure, appointment processes, and culture (Purohit and Martineau, 2016; Purohit et al., 2014, 2016). To inform our thinking during this formative phase, we conducted qualitative interviews with district public health managers. Previous studies on management practices provided the starting point for developing a topic guide that explored the relevance of different dimensions of management, the language used by public health managers to describe their duties and roles, and the management systems of government. We conducted 12 in-depth interviews with public health managers in four districts in Maharashtra. Understanding of government systems and processes was critical. To give one example, each health facility has a patient welfare committee, known as rogi kalyan samite (RKS), which is permitted to generate and manage funds locally. Frequent use of RKS funds for the local procurement of drugs is indicative of a poorly managed district drug supply chain.

These in-depth interviews as well as tools used previously in the literature informed the development of a first version of the management survey tool that contained questions on management practices structured around five broader dimensions of management found to be relevant: operations; performance monitoring; targets; people management; and autonomy. Multiple iterations of the tool were piloted with district public health managers in five districts of Gujarat to scrutinise further the relevance of each management practice, refine language, and develop a set of probing questions.

3.2.2. Management practices and scoring grid

The tool defined sixteen management practices, grouped into the five management dimensions, as shown in Fig. 1. The complete tool is available in Appendix 2. The general approach of the tool followed Bloom and Van Reenen (2007). Under each management practice, we asked a series of open-ended questions that required the respondent to elaborate beyond a simple yes or no answer, making the interview feel more like a conversation and helping respondents to be more at ease. Responses to the open ended questions provided the basis to quantitatively score the management practice between 1 (worst) and 5 (best). We evaluated each management practice with three questions such that a total of 48 responses were scored across the sixteen management practices. Table A1 in Appendix illustrates with six example

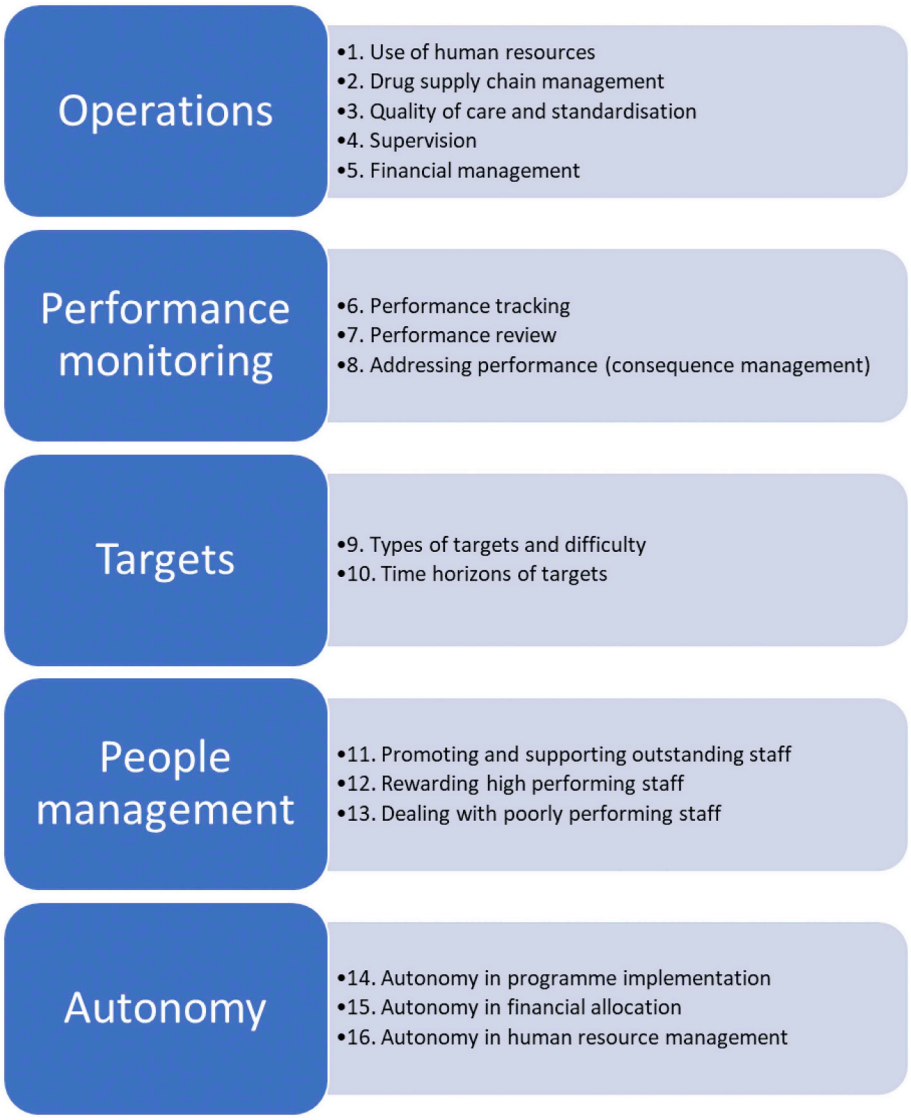

Fig. 1. Management practices and broader dimensions of management.

management practices what the tool was seeking to test and the questions asked in interview.

The scoring options were defined using descriptors to guide interviewers towards an objective assessment of the management practice (see Appendix 2). This was one area in which the tailoring of the tool to the study setting was crucial. The definition of best practice needed to be applicable to district health administration in India and it had to be plausibly obtainable. We were concerned that had we used international norms of best practice in private sector firms, we would have recorded low scores with little or no variation in the quality of management across districts.

3.3. Data collection

3.3.1. Type of interview and respondents

We considered gathering data in a group interview with several managers in an organisation but dismissed this option because of the strong hierarchical relationships in the Indian government and the reluctance of junior staff to speak up in the presence of more senior colleagues. Instead we conducted face-to-face interviews with individuals. These were undertaken in private and respondents were assured of confidentiality of their responses through the informed consent procedure. Our judgement was that telephone interviews were not a viable option – they would have resulted in refusals and inaccurate responses.

The study was conducted in 34 districts of Maharashtra between April and July 2016. We did not include Mumbai which is the state capital. We sought to interview up to three different managers within each district office. Interviewing different respondents within each district office allowed us to capture information from someone junior enough to know actual day-to-day practices and someone senior enough to know and understand the broader context. Our piloting suggested that it was important to capture information from managers with different perspectives afforded by their roles since some managers were more or less informed about specific management practices.

Respondents were eligible for interview if they had been employed in the present post (either permanent or acting) for at least three months prior to interview. We approached the following district public health managers, in order of priority: Chief District Health Officer (DHO); Assistant District Health Officer (ADHO); District Program Manager (DPM); Reproductive and Child health Officer (RCHO); and Quality Medical Officer (QMO). If any of the first three were not available for interview, we approached the fourth or fifth on the list.

We obtained the permission and support of the State Principal Secretary of Health to carry out the study. She provided a letter of support which was shown to each respondent when introducing the

study. Teams were given standard guidance in how to schedule interviews and conduct them. In advance of arriving in the district, researchers contacted eligible respondents to schedule interviews. All interviews were voluntary and had a duration of about 1 h. The most challenging aspect of obtaining interviews was securing the availability of the eligible respondent. Most were very busy but once they had committed to being interviewed the interviews ran smoothly.

3.3.2. Interview procedures

Interviews were conducted by field teams composed of two members, a primary interviewer in charge of leading the interview and asking questions, and a second interviewer instructed to take detailed notes throughout the conversation. Interviewers prepared all necessary materials to administer the survey before entering the location of the interview. The primary interviewer explained the purpose of the study and sought consent. The introduction to the study emphasised: confidentiality of the information provided, focus of the interview on actual practices and not on general functioning of government systems, request for honest and frank responses, and encouragement to discuss challenges and experiences faced by respondents.

The interview followed the sequence of questions in the survey tool. The scoring grid allowed the interviewer to score each management practice question on a scale from 1 (worst) to 5 (best). There was no scope for providing “don't know” responses. The range of 1–5 was not intended to be a subjective scale; the scoring guidance provided an objective description of what scores of 1, 3 or 5 meant. Although the guidance did not provide a description for score 2 and 4, interviewers were free to give any score on the 1 to 5 scale according to their best judgment of mangers' responses. Interviewers were instructed to avoid using a score of 3 as a default in case they experienced difficulties in assigning a score. Instead, they were provided with instructions to probe respondents to get enough information to score practices.

In order to facilitate the scoring process, there were two versions of the tool. The first was the tool used during interview that contained a list of questions and a space for notes for each management practice; the second was a scoring version that was used by interviewers to score answers immediately after the end of the interview. This scoring version was kept hidden from the respondent for the duration of the interview. Both team members were responsible for scoring management practices at the end of the interview. The scoring guidelines were used to discuss and find agreement on scoring. In the early stage of the fieldwork, teams were sometimes assisted by a researcher, responsible for taking extensive notes, helping with facilitation, and advising interviewers if inaccurate scoring was identified.

3.3.3. Limiting survey bias

We employed a number of well-tested strategies to limit survey bias during data collection (Bloom and Van Reenen, 2010b). On the respondent side, we interviewed up to three respondents per district, to limit the influence of any single respondent and reduce the amount of noise in the district level measure of management. Respondents were blinded to the scoring to limit bias that might arise if they had the impression they were being assessed. They were not informed their responses would be scored and the scoring was done confidentially out of sight.

Piloting of the tool suggested that respondents, particularly those in the most senior position, had a tendency to describe government management systems and policies that existed on paper when it was actual practices that we were after. This is similar to what has been referred to as the “public” and “private” face of individuals – the former representing how people present their views to strangers, the latter representing what people divulge to trusted friends (Goffman, 1959). To get beyond the public face, interviewers were trained to use techniques to obtain more detail on actual experiences and practices in instances where the original questions did not elicit the necessary information. These included probing, asking for examples, and asking for direct personal experience to steer the conversation towards actual practices and strategies adopted (see Appendix 4).

To limit interviewer bias, the scoring of management practices was based on exact descriptors across the range of scores in order to reduce the role for subjective interpretation. We hired interviewers with good knowledge of management in the district health administration and conducted intense training and mock interviews over a one week period to calibrate scores between different interviewers as a means to improve consistency. Finally, both interviewers were responsible for scoring. After the interview had been completed, the interviewers would refer to their notes and agree on a score for each question. Where an individual score could not be agreed, further input was sought from a member of the core research team.

3.4. Psychometric performance

We undertook a range of approaches to assess the acceptability, reliability and validity of the management measurement tool. We were guided by the framework presented in Smith et al. (2005) that recommends a number of commonly used psychometric tests to determine whether a measurement tool provides scientifically credible information. Tests were performed on individual items as well as the summary score, calculated as an unweighted average of the responses to all the questions, hence scaled between 1 and 5. Assessment of the psychometric performance of the tool was done ex post, not to inform the development of the scale.

First, we carried out item analysis tests and assessed acceptability of the tool. Item analysis identifies questions (items) that have weak psychometric performance based on the following tests: unrotated principal component factor analysis to determine whether items are measuring a single factor; the extent of missing data; maximum endorsement frequencies as indicated by the proportion of respondents endorsing each response category; floor and ceiling effects as indicated by the proportion of respondents endorsing the maximum and minimum response categories; the extent of item redundancy as indicated by inter-item correlations; and internal consistency as measured by item-total correlations. Acceptability refers to the quality of the data in terms of completeness and score distributions. We examined the extent of missing data as well as ceiling and floor effects in the summary score. Table 2 summarises the tests and criteria used.

Second, we examined the reliability of the tool. Reliability concerns the internal consistency and test-retest reliability of the data. We assessed the former – the extent to which items comprising a scale measure the same construct – using Cronbach's alpha. We assessed the latter by measuring the within district variation between respondents with the intraclass correlation. An analysis of variance model is used to estimate the intraclass correlation. Because this test is performed on data from *different* respondents within the same district, it is not intended as a pure test-retest reliability measure. However, it remains highly informative given that the purpose of the tool was to capture management practices at the district level.

We further investigated the reliability of the tool by examining whether there were systematic differences in the overall management score between respondents within the same district. Specifically, we analysed whether characteristics of the respondent were associated with the overall management score using a linear mixed effects model that allowed for district random effects. Characteristics included were the job title of the respondent, whether the position was permanent, whether the appointment was through an internal promotion or external process, gender, tenure in the current post (years), and tenure working for the government in the district (years).

Third, we present evidence on the construct validity of the tool. We examined whether management practices were correlated with the number of staff in the district health office with a management qualification. We defined a management qualification as either a Master of Business Administration (MBA) or a Post Graduate Diploma in Public

Health Management (PGDPHM). We hypothesised that, conditional on the total number of staff, districts with a higher number of staff with a management qualification would have better management scores. Using the same linear mixed effects model as previous, we ran a regression of the summary management score on the number of district staff with a management qualification and the total number of district staff. We included controls for characteristics of the respondent. We also conducted exploratory factor analysis on the 16 management practices to assess the importance of individual practices and to determine the extent of support for subscales.

It is important to note that we used responses to all the questions when generating an overall measure of management. In other words, the item analysis tests described previously were used to make an assessment of the psychometric performance of individual items, not to eliminate items in the development of the overall score. We do, however, examine the sensitivity of our results to using an overall management score that is based only on those questions that performed strongly in the item analysis.

4. Results

4.1. Descriptive statistics

We conducted 99 interviews in the 34 study districts. We interviewed the target number of three respondents in 31 (91%) districts, and two respondents in 3 (9%) districts. Table 1 shows the characteristics of the respondents. They were on average 40 years of age and had been in post for three years while working in the district government for almost five years. Most respondents (87%) were men. The majority of interviews were with the three most senior managers in each district; in six districts the Chief District Health Officer was not available for interview.

Almost two-fifths of respondents (39%) were in a permanent position, with the remaining 60% of respondents temporarily covering a position higher than their current grade (acting) or working in their position under a fixed term contract. Most appointments (78%) were made directly, which means that officials were recruited through an external process to fill a specific position rather than being promoted internally based on seniority and eligibility by the Departmental Promotion Committee (DPC). The most common highest qualification was Doctor of Medicine (MD), followed by Master of Business Administration (MBA) and Diploma in Public Health (DPH).

Table 1
Descriptive statistics on respondent characteristics.

Variable	n/N (%) or mean (SD)
Age (years)	40.2 (7.5)
Tenure in post (years)	2.8 (2.5)
Tenure working in district (years)	4.6 (5.2)
Gender	
Male	86/99 (87%)
Female	13/99 (13%)
Position	
Chief District Health Officer (DHO)	28/99 (28%)
Additional District Health Officer (ADHO)	23/99 (23%)
District Program Manager (DPM)	32/99 (32%)
District Reproductive and Child Health Officer (DRCHO)	14/99 (14%)
District Surveillance Officer (DSO)	2/99 (2%)
Type of position	
Permanent	39/99 (39%)
Acting position	28/99 (28%)
Contract	32/99 (32%)
Appointment	
Direct	77/99 (78%)
Departmental Promotion Committee	22/99 (22%)
Highest degree	
MD	41/99 (41%)
Master of Business Administration	19/99 (19%)
Diploma in Public Health	15/99 (15%)
PhD	2/99 (2%)
Master's in Public Health	5/99 (5%)
Other	17/99 (17%)

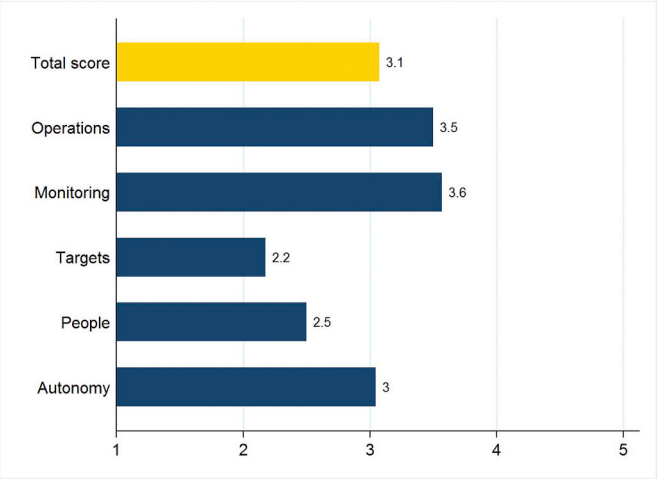

Fig. 2. Management practice score by dimension.

The mean overall management score was 3.1 (see Fig. 2). Districts scored highest on monitoring (3.6), followed by operations (3.5), autonomy (3.0), people (2.5) and targets (2.2), providing an indication of the better performing dimensions of management. The distribution in the overall summary score was reasonably narrow (standard deviation 0.36), with few scores below 2.5 (5 respondents) and above 4.0 (3 respondents).

4.2. Item analysis and acceptability

Table 2 presents the results of the psychometric tests. The first set of results under item analysis identifies questions with poor psychometric performance. The loadings on the first principal component analysis factor ranged from 0.004 to 0.70. Of the 48 items, 16 items failed to load more than 0.3 on the first principal component analysis factor. The inter-item correlation was less than 0.75 for all items such that no item failed the item redundancy test. The item-total correlation was less than 0.25 for 15 out of 48 items. No item failed the maximum endorsement frequency test and no item had more than 5% missing data. There were no observations with missing data and there was a reasonably even distribution in the score. There were no floor or ceiling effects in that no observation had the minimum value of zero or the maximum value of five.

4.3. Reliability and within district variation

The summary management score showed good internal consistency. Cronbach's alpha for the overall score was 0.904, well above the standard threshold of 0.7. The district intraclass correlation was 0.52, indicating that more than 50% of the variation in management practices was between districts. This suggests reasonable agreement between respondents and test-retest reliability, and gives us confidence that the reported scores contain a strong signal of actual management practice. Measurement error is nonetheless an important issue. We presume that much of the 48% of the variation in management practices that is due to differences between respondents within the same districts reflects survey measurement error.

Table 3 reports the coefficients and residual intraclass correlations from three models examining how the overall score differed between respondents in the same district. In the first model, with district random effects only, the intraclass correlation was similar to what was reported

Table 2
Psychometric tests.

Psychometric property	Criteria	Result	Failed items
Item analysis	All items should load on first principal component analysis factor > 0.3	16/48 failed	5c, 6b, 8a, 8c, 9a, 9b, 10a, 10c, 11b, 11c, 12a, 12b, 13a, 16a, 16b, 16c
Item analysis	Inter-item correlation should be < 0.75	No item failed	
Item analysis	Item-total correlation should be > 0.25	15/48 failed	5c, 6b, 8a, 9a, 9b, 10a, 10c, 11b, 11c, 12a, 12b, 13a, 16a, 16b, 16c
Item analysis	Maximum endorsement frequency (MEF) should be < 80% (includes floor and ceiling effect < 80%)	No item failed	
Item analysis	Missing data should be < 5%	No item failed	
Acceptability	Missing data of summary score should be < 5%	No missing observations	
Acceptability	Floor and ceiling effect of summary score < 10%	% floor: 0 observations % ceiling: 0 observations	
Reliability	Cronbach's alpha for summary score > 0.7	0.9040	
Reliability	District intraclass correlation for summary score	0.5215	

Table 3
Management score and respondent characteristics.

	Model 1		Model 2		Model 3	
	Coefficient	p value	Coefficient	p value	Coefficient	p value
Chief District Health Officer (CDHO)			0.236	< 0.001	0.244	< 0.001
District Program Manager (DPM)			−0.003	0.955	0.072	0.423
District Reproductive and Child Health Officer (DRCHO)			0.066	0.410	0.131	0.097
District Surveillance Officer (DSO)			−0.242	0.188	−0.076	0.668
Permanent position					0.161	0.040
Departmental promotion committee appointment					0.194	0.018
Age					−0.003	0.484
Male					−0.172	0.022
Tenure in post					0.021	0.042
Tenure in district					−0.0097	0.102
Residual intraclass correlation	0.51		0.63		0.65	
Districts	34		34		34	
Observations	99		99		99	

previously. The second model included the position of the respondent as covariates and indicates systematic differences in management practices according to the position of the respondent. Management practice scores based on interviews with Chief DHOs were significantly higher (equivalent to 0.65 standard deviations) than those from interviews with Additional DHOs. The positions of other respondents were not associated with the management score. The third model included additional characteristics of the respondents and their job. The coefficient on Chief DHO remained positive and statistically significant. Whether the respondent had a permanent position and was appointed through the internal promotions process were significantly associated with better reported management practices. Gender was also significant, with male respondents associated with worse reported management practices. Finally, tenure in the current post was associated with a higher management score. The results remained largely unchanged when we used an overall score of management based on the reduced set of 32 items that survived the item analysis tests (Table A2 to A4 in Appendix) or an overall score based on the primary factor from

Table 4
Management score and management qualifications.

	Model 1		Model 2	
	Coefficient	p value	Coefficient	p value
Number of staff in district health office	0.0048	0.289	0.0057	0.166
Number of staff with a management qualification	0.123	0.002	0.136	< 0.001
Chief District Health Officer (CDHO)			0.239	< 0.001
District Program Manager (DPM)			0.100	0.258
District Reproductive and Child Health Officer (DRCHO)			0.164	0.036
District Surveillance Officer (DSO)			−0.100	0.566
Permanent position			0.173	0.024
Departmental promotion committee appointment			0.230	0.004
Age			−0.002	0.633
Male			−0.164	0.027
Tenure in post			0.024	0.018
Tenure in district			−0.012	0.051
Residual intraclass correlation	0.40		0.51	
Districts	34		34	
Observations	99		99	

factor analysis (Table A5 to A6 in Appendix).

Given that our tool sought to measure management practices at the district level, the results in Table 3 suggest that some of the measurement error was systematic. As shown by the residual intraclass correlations, as we accounted for more respondent characteristics, the share of the variation in management practices driven by differences across respondents within districts decreased.

4.4. Validity

Table 4 reports the results showing that the number of staff with a management qualification was positively associated with better management practices. In the first model that allowed for random effects at the district level, adding one additional member of staff with a management qualification in the district health office was associated with an increase of 0.12 in the management score ($p = 0.002$). In the second model, the coefficient on the number of staff with a management qualifications remains strongly positive and highly significant ($p < 0.001$) when respondent characteristics are included. Similar results were obtained when used alternative approaches to the overall score of management (Table A2 to A6 in Appendix).

Exploratory factor analysis on the 16 management practices showed that one principal component loaded positively on all the practices, explaining 36% of the variance (Table A7 in Appendix). This suggests that there is a common factor of "good management" (Bloom and Van Reenen, 2007). DHOs that perform well on one management practice tend to perform well on all management practices. A second principal component accounted for a further 10% of the variance, but the pattern in the factor loadings is difficult to interpret and conceptually unclear.

5. Discussion

In this paper we described the development of a tool to measure management practices in India's district public health administration. The process was systematic, informed by a conceptualisation of what good management looks like in this specific context, qualitative interviews with district health managers, and extensive piloting of the tool. We then collected data in every district of Maharashtra, interviewing up to three district health managers in each district. The data collection methods were carefully tailored to the study context and documented in detail. Finally, we assessed the acceptability, reliability and validity of the tool.

We discuss in turn a number of key findings. First, it was feasible to implement the tool. The response rate was very high and there were almost no missing data. However, this should not obscure the fact that such research, from a practical perspective, was challenging. It required close engagement with and considerable buy-in from government. Face-to-face interviews were the only feasible option and getting the time of busy public health managers required patience. Our experience also suggested that there is balance to be struck when hiring interviewers. On the one hand, we wanted interviewers with sufficient experience and knowledge of the district health system who could be credible in the eyes of interviewees. On the other hand, we did not want interviewers to personally know the public health managers working there as to generate bias.

Second, the results from the item analysis indicated that overall psychometric performance of the tool was reasonably strong although some items were identified as being redundant. One third of the items failed to load more than 0.3 on the first principal component analysis factor. These items could be regarded as candidates for elimination in the development of the overall summary score of management practices. While our subsequent analyses were based on a summary score that used all the items, we showed that the results were not sensitive to a score based on the reduced set of items. Decisions regarding the retention and elimination of items should give consideration to content validity, and specifically the trade-off between adhering to the conceptual framework and better psychometric properties arising from item reduction (Smith et al., 2005). To the best of our knowledge, this study is the first to report on the item response properties of a scale used to measure management practices.

Third, the results showed that the reliability of the tool was commensurate with others used in the literature and there was evidence of validity. Internal consistency of the tool was found to be high. The ICC of 0.52 showed reasonable test-retest reliability. In the WMS, a second interviewer was used to interview a second plant manager in the same firm showing that the correlation between the two interviews was 0.51 (Bloom and Van Reenen, 2010a). Using a modified management practices tool, another study reports a correlation of 0.55 in the management scores from two independent respondents in the same firm (Bloom et al., 2016a). In a study of business practices, the correlation in scores measured twice in the same firms over a one year period was 0.59 (McKenzie and Woodruff, 2016). While none of these test-retest reliability measures are ideal, they serve to place our results in context.

By examining within district variation in the management score, our study goes beyond the literature in being able to identify potential sources of measurement error. Most notably, the position of the respondent seemed to matter. Management scores from the Chief DHO were systematically higher than other respondents. Multiple interpretations of these data are possible but we believe that Chief DHOs responded by describing the management practices that existed on paper, despite our best efforts to push them towards describing actual practices in place. In other words, their responses were overly optimistic, driven by social desirability bias. The implications of these findings are twofold. It is important to interview respondents who are not so senior as to be unaware of (or unwilling to report) actual day-to-day practices. Respondent characteristics should be included as noise controls in further uses of the data to help remove some of the measurement error in the management score.

With regards to validity, we found that management practices were strongly correlated with the number of staff in the district health office with a management qualification. These findings are consistent with those of Bloom and Van Reenen (2007), who also found that firms with higher skilled staff, as proxied by management qualifications, had better management practices. Factor analysis showed that one principal component loaded positively on all the management practices although there was little support for management sub-scales.

The study had a number of limitations. The tool did not seek to measure leadership practices which, for example, have been shown to be important for staff satisfaction amongst nurses (Cummings et al., 2010). On a related note it is interesting that the leadership literature underscores the point that task-oriented styles are associated with worse performance than relational styles (Bandiera et al., 2017; Cleary et al., 2018; Cummings et al., 2010). There is likely a trade-off between having a tool that is generalisable and one that is sufficiently tailored to the context as to be reliable and valid. We developed the tool for the purposes of measuring management practices in Maharashtra. We believe that the tool could be used to measure management practices in the district health offices of other Indian states with minimal adaptations. Much more work would need to be done to adapt the tool to other countries, although the general framework could be maintained. The validity of the tool was only touched upon. We discuss below future research that could better assess the validity of tool but note that more extensive validation is challenging because there is no gold standard measure of management to assess criterion validity.

There are a number of directions in which we intend to take this research. Future analysis will seek to examine associations between district health management practices and health service coverage in the population by combining these data with large representative household datasets. Findings from such research will provide novel evidence on the question of whether district management matters for population service coverage. Other directions include further work to validate the management practices tool by examining known group differences and

associations with measures of organisational performance. In contrast to private firms in which performance is easily measured by profitability, survival, and market value, public sector organisational performance is much harder to gauge and alternative measures must be sought, such as project completion and budget execution rates. We envisage expanding data collection on district health management practices to other states in India, incorporating questions on leadership and eventually using the tool to test the effectiveness of management strengthening interventions.

Acknowledgements

We gratefully acknowledge the advice and support given by Dutte Meghe, Daniela Scur and Sarah Smith. The research was funded by the Medical Research Council, the Economic and Social Research Council, the UK's Department for International Development and the Wellcome Trust under the Health Systems Research Initiative (MR/M002179/1).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2018.10.031.

References

Bandiera, O., Hansen, S., Prat, A., Sadun, R., 2017. CEO Behaviour and Firm Performance. LSE, London.
Bloom, N., Van Reenen, J., 2007. Measuring and explaining management practices across firms and countries. Q. J. Econ. 122 (4), 1351–1408.
Bloom, N., Van Reenen, J., 2010a. Why do management practices differ across firms and countries? J. Econ. Perspect. 24 (1), 203–224.
Bloom, N., Van Reenen, J., 2010b. New approaches to surveying organizations. Am. Econ. Rev. 100 (2), 105–109.
Bloom, N., Lemos, R., Sadun, R., Scur, D., Van Reenen, J., 2014. JEEA-FBBVA lecture 2013: the new empirical economics of management. J. Eur. Econ. Assoc. 12 (4), 835–876.
Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I., Van Reenen, J., 2016a. What Drives Differences in Management?.
Bloom, N., Lemos, R., Sadun, R., Scur, D., Van Reenen, J., 2016b. International data on measuring management practices. Am. Econ. Rev. 106 (5), 152–156.
Bossert, T., 1998. Analyzing the decentralization of health systems in developing countries: decision space, innovation and performance. Soc. Sci. Med. 47 (10), 1513–1527.
Bruhn, Miriam, Karlan, Dean, Schoar, Antoinette, 2018. The impact of consulting services on small and medium enterprises: evidence from a randomized trial in Mexico. J. Polit. Econ. 126 (2), 635–687.
Cleary, S., Toit, A.D., Scott, V., Gilson, L., 2018. Enabling relational leadership in primary healthcare settings: lessons from the DIALHS collaboration. Health Pol. Plann. 33 (Suppl. l_2), ii65–ii74.
Cleary, S.M., Molyneux, S., Gilson, L., 2013. Resources, attitudes and culture: an understanding of the factors that influence the functioning of accountability mechanisms in primary health care settings. BMC Health Serv. Res. 13, 320.
Conn, C.P., Jenkins, P., Touray, S.O., 1996. Strengthening health management: experience of district teams in the Gambia. Health Pol. Plann. 11 (1), 64–71.
Cummings, G.G., MacGregor, T., Davey, M., Lee, H., Wong, C.A., Lo, E., Muise, M., Stafford, E., 2010. Leadership styles and outcome patterns for the nursing workforce and work environment: a systematic review. Int. J. Nurs. Stud. 47 (3), 363–385.
Daire, J., Gilson, L., Cleary, S., 2014. Developing Leadership and Management Competencies in Low and Middle-income Country Health Systems: a Review of the Literature, RESYST Working Paper. LSHTM, London.
de Savigny, D., Adam, T., 2009. Systems Thinking for Health Systems Strengthening. World Health Organization, Geneva.
Dorros, G., 2006. Building Management Capacity to Rapidly Scale up Health Services and Outcomes. World Health Organization, Geneva.
Egger, D., Travis, P., Dovlo, D., Hawken, L., 2005. Strengthening Management in Low Income Countries, Making Health Systems Work. World Health Organisation, Geneva.
Gilson, L., Erasmus, E., 2004. Values in Use and Organisational Culture: Exploring the Relevance to Health Systems Development, UN Millennium Project Task Force on Child Health and Maternal Health.
Gilson, L., Barasa, E., Nxumalo, N., Cleary, S., Goudge, J., Molyneux, S., Tsofa, B., Lehmann, U., 2017. Everyday resilience in district health systems: emerging insights from the front lines in Kenya and South Africa. BMJ Glob Health 2 (2), e000224.
Goffman, E., 1959. The Presentation of Self in Everyday Life Garden. Doubleday, City, NY.
Goldfinch, S., Derouen, K., Pospieszna, P., 2012. Flying blind? Evidence for good governance public management reform agendas, implementation and outcomes in low income countries. Publ. Adm. Dev. 33, 50–61.
Grous, A., 2011. Management Practices in the UK Aerospace Sector, LSE Working Paper.
Hood, C., 2006. Gaming in targetworld: the targets approach to managing british public services. Publ. Adm. Rev. 66 (4), 515–521.
International Institute for Population Sciences (IIPS), & Macro International, 2016. National Family Health Survey (NFHS-4), 2015-16: India: Women's Questionnaire. International Institute for Population Sciences, Mumbai.
Kotter, J.P., 2001. What leaders really do. Harv. Bus. Rev. 79, 85–96.
Lemos, R., Scur, D., 2016. Developing Management: an Expanded Evaluation Tool for Developing Countries. London School of Economics, Centre for Economic Performance, London.
Mangham, L.J., Hanson, K., 2010. Scaling up in international health: what are the key issues? Health Pol. Plann. 25 (2), 85–96.
McCormack, J., Propper, C., Smith, S., 2014. Herding cats? Management and university performance. Econ. J. 124 (578), F534–F564.
McKenzie, D., Woodruff, C., 2016. Business practices in small firms in developing countries. Manag. Sci. 63 (9), 2967–2981.
Mills, A., Bennett, S., Russell, S., 2001. The Challenge of Health Sector Reform. What Must Governments Do? Palgrave, Basingstoke.
Minogue, M., Polidano, C., Hulme, D., 2000. Introduction: the analysis of public management and governance. In: Minogue, M., Polidano, C., Hulme, D. (Eds.), Beyond the New Public Management. Changing Ideas and Practices in Governance. Edward Elgar, Cheltenham, pp. 1–14.
MSH, 2012. Human Resource Management Rapid Assessment Tool for Health Organizations: a Guide for Strengthening HRM Systems. Management Sciences for Health, Cambridge.
Muralidharan, K., Chaudhury, N., Hammer, J., Kremer, M., Rogers, F.H., 2011. Is There a Doctor in the House? Medical Worker Absence in India.
Mutale, W., Vardoy-Mutale, A.T., Kachemba, A., Mukendi, R., Clarke, K., Mulenga, D., 2017. Leadership and management training as a catalyst to health system strengthening in low-income settings: evidence from implementation of the Zambia Management and Leadership course for district health managers in Zambia. PloS One 12 (7), e0174536.
Office of the Registrar General, 2016. Special Bulletin on Maternal Mortality in India 2011-13. Government of India, New Delhi, India.
Office of the Registrar General & Census Commissioner, 2013. Census of India 2011. Government of India, New Delhi, India.
Pritchett, L., 2009. Is India a Flailing State? Detours on the Four Lane Highway to Modernization. Harvard University John F. Kennedy School of Government, Working Paper Series.
Purohit, B., Patel, D., Purohit, S., 2014. A study of organizational values in government run primary health centres in India. J. Health Manag. 16 (2), 301–311.
Purohit, B., Martineau, T., 2016. Initial posting-a critical stage in the employment cycle: lessons from the experience of government doctors in Gujarat, India. Hum. Resour. Health 14 (1), 41.
Purohit, B., Martineau, T., Sheikh, K., 2016. Opening the black box of transfer systems in public sector health services in a Western state in India. BMC Health Serv. Res. 16 (1), 419.
Rasul, I., Rogger, D., 2018. Management of bureaucrats and public service delivery: evidence from the nigerian civil service. Econ. J. 128, 413–446.
Rose-Ackerman, S., 1986. Reforming public bureaucracy through economic incentives? J. Law Econ. Organ. 2 (1), 131–161.
Scott, T., Mannion, R., Davies, H., Marshall, M., 2003. The quantitative measurement of organizational culture in health care: a review of the available instruments. Health Serv. Res. 38 (3), 923–945.
Seims, L.R., Alegre, J.C., Murei, L., Bragar, J., Thatte, N., Kibunga, P., Cheburet, S., 2012. Strengthening management and leadership practices to increase health-service delivery in Kenya: an evidence-based approach. Hum. Resour. Health 10, 25.
Sharma, Smriti, Tarp, Finn, 2018. Does managerial personality matter? Evidence from firms in Vietnam. J. Econ. Behav. Organ. 150, 432–445.
Sheikh, K., Gilson, L., Agyepong, I.A., Hanson, K., Ssengooba, F., Bennett, S., 2011. Building the field of health policy and systems research: framing the questions. PLoS Med. 8 (8), e1001073.
Smith, S.C., Lamping, D.L., Banerjee, S., Harwood, R., Foley, B., Smith, P., Cook, J.C., Murray, J., Prince, M., Levin, E., Mann, A., Knapp, M., 2005. Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. Health Technol. Assess. 9 (10), 1–93 iii-iv.
Travis, P., Bennett, S., Haines, A., Pang, T., Bhutta, Z., Hyder, A.A., Pielemeier, N.R., Mills, A., Evans, T., 2004. Overcoming health-systems constraints to achieve the millennium development goals. Lancet 364 (9437), 900–906.
Vriesendorp, P., De La Peza, L., Perry, C., 2010. Health Systems in Action: an Ehandbook for Leaders and Managers. Management Sciences for Health, Cambridge.
WHO, 2007a. Towards Better Leadership and Management in Health, Making Health Systems Work. World Health Organisation, Geneva.
WHO, 2007b. Everybody's Business. Strengthening Health Systems to Improve Health Outcomes: WHO's Framework for Action. World Health Organization, Geneva.