

Measuring and managing progress in the establishment of basic health services: the Afghanistan Health Sector Balanced Scorecard

Peter M. Hansen^{1*}, David H. Peters¹, Haseebullah Niayesh², Lakhwinder P. Singh³, Vikas Dwivedi³ and Gilbert Burnham¹

¹*Department of International Health, Johns Hopkins Bloomberg School of Public Health, MD, USA*

²*Monitoring and Evaluation Department, Ministry of Public Health, Islamic Republic of Afghanistan*

³*Indian Institute of Health Management Research, Jaipur, India*

SUMMARY

The Ministry of Public Health (MOPH) of Afghanistan has adopted the Balanced Scorecard (BSC) as a tool to measure and manage performance in delivery of a Basic Package of Health Services. Based on results from the 2004 baseline round, the MOPH identified eight of the 29 indicators on the BSC as priority areas for improvement. Like the 2004 round, the 2005 and 2006 BSCs involved a random selection of more than 600 health facilities, 1700 health workers and 5800 patient-provider interactions. The 2005 and 2006 BSCs demonstrated substantial improvements in all eight of the priority areas compared to 2004 baseline levels, with increases in median provincial scores for presence of active village health councils, availability of essential drugs, functional laboratories, provider knowledge, health worker training, use of clinical guidelines, monitoring of tuberculosis treatment, and provision of delivery care. For three of the priority indicators—drug availability, health worker training and provider knowledge—scores remained unchanged or decreased between 2005 and 2006. This highlights the need to ensure that early gains achieved in establishment of health services in Afghanistan are maintained over time. The use of a coherent and balanced monitoring framework to identify priority areas for improvement and measure performance over time reflects an objectives-based approach to management of health services that is proving to be effective in a difficult environment. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: performance measurement; quality; health services management; balanced scorecard; Afghanistan

* Correspondence to: P. M. Hansen, Department of International Health, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe St, Suite E8132 Baltimore, MD 21205, USA. E-mail: phansen@jhsph.edu

INTRODUCTION

Following the fall of the Taliban regime in late 2001, it became clear that health indicators and health services were in a poor state in Afghanistan (MSH, 2002; Waldman and Hanif, 2002; Bartlett *et al.*, 2005; UNICEF, 2006). The Ministry of Public Health (MOPH), faced with a mostly dysfunctional health system, developed a strategy to rapidly expand the geographic scope of a Basic Package of Health Services (BPHS) to meet the health needs of its population, especially addressing rural areas and the needs of women and children (MOPH, 2003). The strategy involves maintaining government focus on stewardship of the health sector, while contracting out service delivery to non-governmental organizations (NGOs) in much of the country, and using a coherent and balanced national monitoring and evaluation framework to measure performance in the delivery of the BPHS. The BPHS is delivered at three types of facilities: Basic Health Centers (BHCs), Comprehensive Health Centers (CHCs) and District Hospitals (DHs).

Loevinsohn and Harding (2005) and Palmer *et al.* (2006) argue that one advantage of contracting health service delivery is the strong focus it places on attainment of measurable results. Within Afghanistan, the MOPH has developed a Balanced Scorecard (BSC) to monitor performance in BPHS delivery and manage contracts with service providers (Peters *et al.*, 2007). Although the BSC has been used previously in a range of healthcare settings (Kaplan and Norton, 1996), the innovative use of the BSC in Afghanistan at the level of a national health system is the first for a developing country (Peters *et al.*, 2007).

The six domains and 29 indicators used in the BSC were developed through a participatory process involving MOPH leadership and technical staff, NGOs and donor and technical agencies, and implemented by independent agencies in collaboration with the MOPH (Peters *et al.*, 2007). The BSC domains summarize performance from the following six perspectives:

- Domain A: Patients and Community
- Domain B: Staff
- Domain C: Capacity for Service Provision
- Domain D: Service Provision
- Domain E: Financial Systems
- Domain F: Overall Vision

After seeing the results from the 2004 BSC, the MOPH identified eight priority areas in need of improvement, based on two main factors: the unsatisfactory level of performance for each indicator and its importance to the MOPH's strategy to improve people's health. These priority areas, listed in the order in which they appear on the BSC, include: establishment of functional village health councils, drug availability, laboratory functionality, provider knowledge, health worker training, availability of clinical guidelines, monitoring of tuberculosis (TB) treatment and provision of delivery care.

The establishment of functional village health councils is part of the MOPH's strategy to make health services more responsive to the needs of communities (MOPH, 2005). Improving the supply of essential drugs and increasing laboratory

capacity to conduct diagnostic tests are part of the MOPH's strategy to ensure that BPHS facilities are prepared to meet the needs of clients. Health worker knowledge, the participation of health workers in training programs and the use of clinical guidelines are necessary for improving quality of care. Monitoring of TB treatment is necessary to implement DOTS, which is required at BHCs, CHCs and DHs. Provision of delivery care through BPHS facilities is part of the MOPH's strategy to make motherhood safer.

The objective of this paper is to use the results of the 2005 and 2006 BSCs to assess whether progress has been made in the priority areas identified from the 2004 BSC. It is hypothesized that increases in performance are greater among the eight priority indicators than among the remaining 21 indicators of the BSC.

METHODS

The data used to calculate BSC indicators are collected by the MOPH each year from July to October. The random sampling of facilities, health workers and patients in each province has been described previously (Peters *et al.*, 2007), and is conducted independently from one round to the next. Table 1 shows the sample size of the first three rounds of the BSC.

Scores were calculated in Stata 8 (Stata Corp, 2004), with the scores for all observations within a province averaged to derive an unadjusted mean provincial score for each indicator. Scores for each indicator were then adjusted according to a standardized distribution of facility type or health worker characteristics, in order to adjust for differences in the distribution of these characteristics between provinces. In order to test for differences in the score for each indicator across different years for the 29 provinces covered in all three rounds, the Wilcoxon matched-pairs signed-rank test, a non-parametric test, was used (Rees 1987). This tests the equality of matched pairs of observations, with the null hypothesis stating that the distribution of provincial scores is the same in the years being compared. A change in indicator values is considered to be statistically significant if the *p*-value is less than or equal to 0.05.

Table 1. Sample size in 2004, 2005 and 2006 BSCs

| | 2004 | 2005 | 2006 |
|--|--------|------|------|
| Provinces | 33 | 30 | 30 |
| Facilities | 617 | 629 | 630 |
| Health worker interviews | 1553 | 1458 | 1723 |
| Community health worker interviews | 167 | 306 | 907 |
| Direct observations of patient-provider interactions | 5719 | 5863 | 5964 |
| Exit interviews | 5597 | 5863 | 5964 |
| Households interviewed | 13 843 | — | — |

Note: In 2004, Daikundi Province was not included in the assessment, since it was newly formed at the time of the assessment and had no functional BPHS facilities. The 2005 and 2006 assessments included Daikundi, but did not include Helmand, Kandahar, Zabul and Uruzgan, since assessment teams could not move safely in those provinces.

Each indicator on the BSC is scored as a percentage or on a scale of 0 to 100. The concentration indices, which measure the distribution of a variable according to level of poverty (Kakwani *et al.*, 1997), were rescaled from its usual range of -1 to 1 to 0 to 100 , with 0 on the concentration index scaled to 50 (perfect equity). For each indicator, upper and lower benchmarks corresponding to the top and bottom quintiles for provincial performance in 2004 were established to indicate levels of performance that are achievable in Afghanistan.

RESULTS

Between 2004 and 2005, seven of the eight priority indicators showed statistically significant improvements, compared to six out of the remaining 21 indicators. The eighth priority indicator, provision of delivery care, showed a statistically significant improvement in 2006 after remaining unchanged between 2004 and 2005. Overall, 17 of the 29 indicators on the 2006 BSC show statistically significant improvements compared to 2004, while one indicator shows a statistically significant decrease (Table 2).¹

Changes in priority indicators

The first priority indicator measures whether facilities have an active village health council and written records of activities conducted by the council. In the province with the median score for this indicator in 2004, only 34.5% of health facilities had active village health councils; in 2005 and 2006, 56.4% ($p=0.05$) and 67.5% ($p<0.01$) of the facilities in the province with the median score had active village health councils, respectively. Between 2004 and 2006, the number of provinces with less than one-third of sampled facilities having written records of health council activities decreased from 15 to 2.

The drug availability index measures the continuous availability of a set of essential drugs at the health facility. In the province with the median score for this indicator in 2004, facilities on average had 68.8% of the items included in the index and four provinces had fewer than 50% of these items. The median provincial score increased to 83.7% by 2005 ($p<0.01$), all provinces met the lower benchmark and over half of the provinces met the upper benchmark. The 2006 median provincial score was approximately the same as that in 2005 (84.6%).

The median provincial score for the laboratory functionality index, which measures CHC and DH capacity to conduct essential diagnostic tests, was 18.8% in 2004, with nine provinces having scores below 10%. Compared to 2004, the median score was higher in 2005 (36.3%, $p<0.01$) and 2006 (43.4%, $p<0.01$). In 2006, all provinces met the lower benchmark and all but five met the upper benchmark.

Between 2004 and 2005, the median provincial score increased for staff receiving in-service training within the last year (39.0–74.5%, $p<0.01$) and provider

¹Scores presented here are limited to the 29 provinces covered in all three rounds of the BSC and thus are slightly different than scores presented elsewhere.

THE AFGHANISTAN HEALTH SECTOR BALANCED SCORECARD

Table 2. BSC results for 29 provinces covered in 2004, 2005 and 2006

| Indicator | Median 2004 | Median 2005 | Median 2006 | <i>p</i> value change 2004–2005 | <i>p</i> value change 2004–2006 |
|---|----------------|----------------|----------------|---------------------------------------|---------------------------------------|
| Domain A: Patients and Community | | | | | |
| 1 Overall patient satisfaction | 84.1 | 86.7 | 86.4 | 0.58 | 0.41 |
| 2 Patient perceptions of quality | 75.4 | 76.4 | 80.3 | 0.24 | 0.02 |
| 3 <i>Written Village Health Council activities in community</i> | <i>34.5</i> | <i>56.4</i> | <i>67.5</i> | <i>0.05</i> | <i><0.01</i> |
| Domain B: Staff | | | | | |
| 4 Health worker satisfaction | 63.5 | 64.3 | 68.1 | 0.16 | <0.01 |
| 5 Salary payments current | 66.3 | 89.8 | 80.8 | <0.01 | 0.16 |
| Domain C: Capacity for Service Provision | | | | | |
| 6 Equipment functionality | 65.7 | 66.4 | 77.6 | 0.39 | <0.01 |
| 7 <i>Drug availability</i> | <i>68.8</i> | <i>83.7</i> | <i>84.6</i> | <i><0.01</i> | <i><0.01</i> |
| 8 Family Planning availability | 61.4 | 65.5 | 83.7 | 0.33 | <0.01 |
| 9 <i>Laboratory functionality (CHCs & DHs)</i> | <i>18.8</i> | <i>36.3</i> | <i>43.4</i> | <i><0.01</i> | <i><0.01</i> |
| 10 Meeting minimum staff guidelines | 38.0 | 57.5 | 66.3 | <0.01 | <0.01 |
| 11 <i>Provider knowledge</i> | <i>54.0</i> | <i>69.3</i> | <i>69.3</i> | <i><0.01</i> | <i><0.01</i> |
| 12 <i>Staff received training in last year</i> | <i>39.0</i> | <i>74.5</i> | <i>68.8</i> | <i><0.01</i> | <i><0.01</i> |
| 13 HMIS use | 69.4 | 68.0 | 77.7 | 0.26 | 0.25 |
| 14 <i>Clinical guidelines</i> | <i>34.7</i> | <i>48.8</i> | <i>62.5</i> | <i><0.01</i> | <i><0.01</i> |
| 15 Infrastructure | 55.0 | 44.3 | 48.6 | <0.01 | 0.14 |
| 16 Patient records | 66.8 | 63.3 | 69.2 | 0.13 | 0.87 |
| 17 <i>Monitoring of TB treatment</i> | <i>15.8</i> | <i>21.7</i> | <i>38.0</i> | <i>0.05</i> | <i><0.01</i> |
| Domain D: Service Provision | | | | | |
| 18 Patient history and physical exam | 70.6 | 72.9 | 82.5 | 0.02 | <0.01 |
| 19 Patient counselling | 33.2 | 35.7 | 37.0 | 0.51 | 0.50 |
| 20 Proper sharps disposal | 62.2 | 51.9 | 76.4 | 0.14 | 0.27 |
| 21 Average new outpatient visits per month (BHC > 750 visits) | 21.1 | 33.3 | 56.3 | 0.02 | <0.01 |
| 22 Time spent with patient (> 9 minutes) | 18.5 | 5.9 | 6.9 | <0.01 | 0.02 |
| 23 Provision of antenatal care | 58.5 | 79.1 | 84.7 | <0.01 | <0.01 |
| 24 <i>Provision of delivery care</i> | <i>25.4</i> | <i>22.2</i> | <i>41.4</i> | <i>0.60</i> | <i><0.01</i> |
| Domain E: Financial Systems | | | | | |
| 25 User fee guidelines | 90.6 | 86.6 | 82.5 | 0.34 | 0.14 |
| 26 Exemptions for poor patients | 83.4 | 92.4 | 100 | 0.07 | 0.06 |
| Domain F: Overall Vision | | | | | |
| 27 Females as percentage of new outpatients | 55.2 | 57.4 | 58.0 | 0.02 | <0.01 |
| 28 Outpatient visit concentration index | 50.5 | 50.6 | 51.2 | 0.71 | 0.82 |
| 29 Patient satisfaction concentration index | 50.0 | 49.8 | 49.8 | 0.21 | 0.30 |

Note: Italics denote that the indicator was prioritized for improvement by the MOPH.

knowledge (54.0–69.3%, $p < 0.01$). Between 2005 and 2006, however, provider knowledge remained unchanged, while staff receiving in-service training decreased (74.5–68.8%, $p = 0.04$).

The clinical guidelines index measures the presence of basic guidelines for health workers. The median provincial score increased from 34.7% in 2004 to 48.8% in 2005 ($p < 0.01$) and 62.5% in 2006 ($p < 0.01$). The number of provinces with scores below 30% decreased from 12 to 0 between 2004 and 2006.

The score for maintaining active TB registers was the lowest among all indicators on the 2004 BSC. The median provincial score increased from 15.8% in 2004 to 21.7% in 2005 ($p = 0.05$) and 38.0% in 2006 ($p < 0.01$). The number of provinces with scores lower than 10% decreased from nine to one between 2004 and 2006.

The delivery care indicator measures whether facilities meet the BPHS requirements for provision of delivery care. BHCs must manage routine deliveries, while CHCs must manage routine deliveries, have the ability to do blood-typing and cross-matching and have a partograph. In addition to the above, DHs must have the ability to provide an emergency caesarian section. In 2005, the median provincial score of 22.2% was not statistically different than the 2004 median score (25.4%) and approximately the same number of provinces met the upper and lower benchmarks each year. Over this time period, gains in delivery care were achieved at the DH level, but not at BHCs and CHCs (results are summarized in Table 3).

The 2006 median score of 41.4% was significantly higher than both the 2004 and 2005 scores ($p < 0.01$). Between 2005 and 2006, gains in delivery care were achieved at all three types of facilities, but BHCs were still less likely than higher level facilities to manage routine deliveries (61.1% vs 86.2% at CHCs and 100% at DHs).

Changes in other indicators

Among the remaining 21 indicators on the BSC, one shows a statistically significant decrease, 9 show statistically significant increases and 11 are not statistically different in 2006 compared to 2004. The only indicator that decreased between 2004 and 2006 was the average time spent with patient (18.5% vs 6.9%, $p < 0.01$).

The other indicators for which the 2006 scores are significantly higher than the 2004 scores include patient perceptions of quality (75.4% vs 80.3%, $p = 0.02$), health worker satisfaction (63.5% vs 68.1%, $p < 0.01$), presence of functional equipment (65.7% vs 77.6%, $p < 0.01$), family planning availability (68.8% vs 84.6%, $p < 0.01$), meeting minimum staffing guidelines (38.0% vs 66.3%, $p < 0.01$), the patient history and physical examination index (70.6% vs 82.5%, $p < 0.01$), average new outpatient visits per month at BHCs (21.1% vs 56.3%, $p < 0.01$), provision of antenatal care (58.5% vs 84.7%, $p < 0.01$) and females as a percentage of new outpatients (55.2% vs 58.0%, $p < 0.01$).

Changes in provincial performance

More provinces achieved the upper and lower benchmarks for individual indicators in 2006 compared to 2004. For the 29 provinces covered in all three rounds, the overall mean score across the 29 indicators on the BSC increased from 52.2 to 63.5

Table 3. Results for components of delivery care indicator

| | 2004 | | | 2005 | | | 2006 | | |
|--|------|------|------|------|------|------|------|------|------|
| | BHC | CHC | DH | BHC | CHC | DH | BHC | CHC | DH |
| Percentage managing routine deliveries | 41.6 | 52.5 | 52.0 | 40.0 | 71.4 | 81.4 | 61.1 | 86.2 | 100 |
| Percentage with functional fetoscope | 43.0 | 53.5 | 52.9 | 51.9 | 81.7 | 86.1 | 81.3 | 91.6 | 97.6 |
| Percentage with partograph | 9.3 | 6.4 | 17.0 | 9.5 | 22.6 | 55.8 | 22.5 | 38.3 | 61.9 |
| Percentage able to do bloodtyping and cross-matching | NA | 7.1 | 18.0 | NA | 12.4 | 55.8 | NA | 17.7 | 73.8 |
| Percentage able to manage emergency caesarian sections | NA | NA | 9.3 | NA | NA | 53.5 | NA | NA | 71.4 |

Note: BHCs are not required by the BPHS to do bloodtyping and cross-matching, and BHCs and CHCs are not required to manage emergency caesarian sections.

between 2004 and 2006. Twenty seven of 29 provinces improved their mean scores over this time period, and more than half the provinces gained more than ten points.

DISCUSSION

The 2005 and 2006 BSCs demonstrate that the MOPH and its development partners have made substantial progress in establishing a nationwide network of functional primary care facilities following decades of conflict and deterioration of the public health infrastructure. The biggest progress was made in the priority areas identified by the MOPH based on the results from the 2004 BSC. Although it is difficult to directly attribute changes in BSC indicators to specific causes, the use of a coherent and balanced monitoring and evaluation framework to identify priority areas for improvement and to measure performance over time reflects an effective objectives-based approach to management of health service delivery.

The 2006 BSC shows that BPHS providers made further improvements for five of the priority indicators, while drug availability and provider knowledge stagnated at 2005 levels and health worker training decreased in 2006 after large gains were achieved in 2005. This may reflect shifting priorities after progress was demonstrated in these areas on the 2005 BSC, or it may reflect the challenge of sustaining early gains in a difficult environment. Further efforts are needed over time to ensure that essential drugs are available at all BPHS facilities and that health workers receive regular in-service training to improve their knowledge and support improvements in the quality of care provided. It is essential that the early gains achieved in establishing health services in the early reconstruction period in Afghanistan are maintained over time, to further improve the capacity and quality of health services in a setting where the health status of the population continues to be among the poorest in the world.

In spite of the gains in several priority indicators, the absolute level of performance in many areas—including provision of delivery care, laboratory capacity and monitoring of TB treatment—remains inadequate, and further improvements are required on a priority basis.

The 2005 BSC indicated an urgent need for improvements in provision of delivery care at lower levels of the health system. Between 2005 and 2006, gains were achieved at lower levels of the system, but BHCs and CHCs continued to lag behind DHs. The larger improvements in provision of delivery care at DHs compared to lower level facilities may reflect more intensive resource inputs at higher levels of the system. It is imperative that further gains in provision of delivery care be achieved at lower levels of the health system, since most people do not have access to district hospitals and referral services between levels of the health system remain weak. The health impact in rural areas is likely to be greater if the capacity and quality of delivery services is improved at peripheral facilities and if home deliveries can be made safer. Health managers, central planners and donors who are making allocation decisions that influence geographic distribution of resources need to determine why BHCs and CHCs have lagged behind DHs in improvements in delivery care, and how success can be achieved at other facility levels. One barrier to improving delivery care is the lack of female staff, especially at the BHC level. Between 2004 and 2006,

increases in the percentage of sampled facilities with at least one female doctor, nurse or midwife occurred at the BHC (35.6–66.0%, $p < 0.01$), CHC (42.8–90.2%, $p < 0.01$) and DH levels (41.2–100%, $p < 0.01$), but further increases are needed in a setting where it is rare for a woman to receive delivery care from a man.

The large decrease and low absolute level for the average time spent per patient raise concerns regarding the level of service quality provided by health workers in Afghanistan. Two process measures of quality based on direct observation of patient-provider interactions—the patient history and physical examination index (indicator 18) and the patient counselling index (indicator 19)—provide further insights on the quality of care provided by health workers in Afghanistan. The median provincial score for the indicator of health worker performance in taking patient histories and conducting physical examinations was 70.6% in 2004 and 82.5% in 2006 ($p < 0.01$), indicating that health workers are performing reasonably well in this area.

In comparison, performance in-patient counselling is much lower, with median provincial scores of 33.2% and 37.0% in 2004 and 2006 ($p = 0.50$), respectively. This poor level of performance in communication suggests a possible lack of perceived need on the part of the health worker, a lack of training in this area or inadequate time when there is pressure to see more patients. The low level of performance in this area is consistent with findings elsewhere that improvement in communication by health workers to patients or caretakers lags behind other performance indicators (Pariyo *et al.*, 2005). The root causes of the large gap between these two indicators and the low level of performance in provision of information and counselling requires further investigation and remedial action by the MOPH and service providers.

The BSC is flexible enough to be used for purposes other than identification of national priorities and provincial comparisons. Additional analyses of the BSC are conducted at the request of the leadership of the MOPH to address policy issues, such as the different contracting mechanisms used to engage NGOs that deliver services. Individual NGOs have requested that results be broken down by NGO, to facilitate identification of areas of strength and weakness relative to other NGOs working in the same and other provinces. These supplementary scorecards reflect a high level of demand for the information presented on the BSC and enable the MOPH to better manage contracts and maintain a strong focus on attainment of measurable results.

The BSC provides an evidence base that empowers the MOPH to hold contracted service providers accountable for their performance. MOPH leadership and technical staff frequently use the BSC in meetings with donors and BPHS implementers. For NGO and Provincial Public Health Office (PPHO) staff working in provinces holding Performance-based Partnership Agreements (PPAs) with the MOPH, financial incentives are given for achievement of performance targets. The MOPH gives a performance bonus worth 1% of the contract value if they achieve an increase of 10 points in their mean score across the 29 indicators on the BSC. Based on their scores on the 2005 BSC, NGOs and PPHO staff working in two provinces received performance bonuses. An international NGO working in one province had its contract terminated for unsatisfactory performance based partially on BSC results, and in early 2006 a new NGO won a competitive bidding process to deliver health

services in the province. When the first round of PPA contracts expired in early 2006 and the MOPH sought to enter into a new round of contracts with NGOs, the BSC was used to assess the desirability of re-engaging each NGO in a new contract to deliver services.

Changes to existing indicators on the BSC will be needed over time to serve the needs of an evolving health sector. Several indicators provide information required by the MOPH at an early stage of development of the health sector in the post-conflict period. It was important for the MOPH and its partners to know, for example, whether females are at a disadvantage compared to males in utilizing health services. Among the 29 provinces, 20 provinces saw more females than males as new outpatients in 2004 and 26 saw more females than males as new outpatients in 2006, indicating that females utilize outpatient services in Afghanistan more frequently than males. Although this indicator is not adjusted for level of need, these results indicate that the concern regarding a possible female disadvantage in utilization of outpatient services compared to males is likely a less serious problem than low levels of utilization of antenatal and delivery care by women. Beyond this, it is not clear what an appropriate target level for this indicator should be. This and other indicators on the BSC may be updated as appropriate.

As a next step, the MOPH should incorporate findings from the BSC into the supervision process, with a view toward supporting improvement in specific priority areas, and develop mechanisms for the transfer of knowledge and country-specific best practices among facilities, provinces and agencies. More in-depth analysis of multiple factors that may affect performance can be conducted on a complementary basis, but it is important to keep the BSC simple and understandable to a broad range of stakeholders and to produce results on a timely basis, so that they can feed into the planning cycle. Interventions that are employed to improve performance in different areas should be evaluated and documented, the results shared with others and the lessons learned applied internally and externally. The BSC has been successful in describing the health sector of Afghanistan, measuring change in performance over time, enabling the MOPH and service providers to improve the quality of health services and making service providers more accountable to catchment area communities, the government and donor agencies.

ACKNOWLEDGEMENTS

This study was funded by a contract between the Afghanistan Ministry of Public Health and the Johns Hopkins University, in collaboration with the Indian Institute of Health Management Research. The contributions of the Third Party Evaluation team and the Ministry of Public Health's Monitoring and Evaluation Department are gratefully acknowledged, along with the suggestions from the editors and the anonymous reviewers.

REFERENCES

Bartlett LA, Mawji S, Whitehead S, *et al.* 2005. Where giving birth is a forecast of death: maternal mortality in four districts of Afghanistan, 1999–2002. *Lancet* **365**: 864–870.

THE AFGHANISTAN HEALTH SECTOR BALANCED SCORECARD

- Kakwani N, Wagstaff A, van Doorslaer E. 1997. Socioeconomic inequalities in health: measurement, computation and statistical inference. *J Econ* **77**: 87–103.
- Kaplan RS, Norton DS. 1996. *The Balanced Scorecard*. Boston: Harvard Business School Press.
- Loevinsohn B, Harding A. 2005. Buying results? Contracting for health service delivery in developing countries. *Lancet* **266**: 676–681.
- Management Sciences for Health. 2002. *Afghanistan National Health Resource Assessment*. Health and Development Services, Management Sciences for Health: Boston.
- Ministry of Health. 2003. *A Basic Package of Health Services for Afghanistan*. Islamic Transitional Government of Afghanistan: Kabul.
- Ministry of Public Health. 2005. *A National Health Policy 2005–2009 and National Health Strategy 2005–2006*. Islamic Republic of Afghanistan: Kabul.
- Palmer N, Strong L, Wali A, Sondorp E. 2006. Contracting out health services in fragile states. *Br Med J* **332**: 718–721.
- Pariyo G, Gouws E, Bryce J, Burnham G. 2005. Improving facility based care for sick children in Uganda: training is not enough. *Health Policy Plan* **20**(S1): i58–i68.
- Peters DH, Noor AA, Singh LP, Kakar FK, Hansen PM, Burnham G. 2007. A balanced scorecard for health services in Afghanistan. *Bull World Health Organ* **85**(2): 146–151.
- Rees DG. 1987. *Foundations of Statistics*. Chapman and Hall: London.
- Stata Corp. 2004. *Stata Statistical Package: Release 8*. Stata Corp LP: College Station, TX.
- UNICEF. 2006. *Best Estimates of Social Indicators for Children in Afghanistan: 1990–2005*. United Nations Children’s Fund: New York.
- Waldman R, Hanif H. 2002. *The Public Health System in Afghanistan: Current Issues*. Afghanistan Research and Evaluation Unit: Kabul.