Rating Maternal and Neonatal Health Programs in Developing Countries

Rodolfo A. Bulatao and John A. Ross

August 2000

Carolina Population Center
University of North Carolina
at Chapel Hill
123 W. Franklin Street
Suite 304
Chapel Hill, NC 27516
Phone: 919-966-7482
Fax: 919-966-2391
measure@unc.edu
www.cpc.unc.edu/measure

Collaborating Partners:

Macro International Inc.
11785 Beltsville Drive
Suite 300
Calverton, MD 20705-3119
Phone: 301-572-0200
Fax: 301-572-0999
measure@macroint.com

John Snow Research and Training Institute
1616 N. Ft. Myer Drive
11th Floor
Arlington, VA 22209
Phone: 703-528-7474
Fax: 703-528-7480
measure_project@jsi.com

Tulane University
1440 Canal Street
Suite 2200
New Orleans, LA 70112
Phone: 504-584-3655
Fax: 504-584-3653
measure2@tulane.edu

Funding Agency:

Center for Population, Health
and Nutrition
U.S. Agency for
International Development
Washington, DC 20523-3600
Phone: 202-712-4959

WP-00-26

The research upon which this paper is based was sponsored by the MEASURE *Evaluation* Project with support from the United States Agency for International Development (USAID) under Contract No. HRN-A-00-97-00018-00.

The working paper series is made possible by support from USAID under the terms of Cooperative Agreement HRN-A-00-97-00018-00. The opinions expressed are those of the authors, and do not necessarily reflect the views of USAID.

The working papers in this series are produced by the MEASURE *Evaluation* Project in order to speed the dissemination of information from research studies. Most working papers currently are under review or are awaiting journal publication at a later date. Reprints of published papers are substituted for preliminary versions as they become available. The working papers are distributed as received from the authors. Adjustments are made to a standard format with no further editing.

A listing and copies of working papers published to date may be obtained from the MEASURE *Evaluation* Project at the address listed on the back cover.

Other MEASURE *Evaluation* Working Papers

WP-00-25 Abortion and Contraceptive Use in Turkey (Pinar Senlet, Jill Mathis, Siân L. Curtis, and Han Raggers)

WP-00-24 Contraceptive Dynamics among the Mayan Population of Guatemala: 1978-1998 (Jane T. Bertrand, Eric Seiber and Gabriela Escudero)

WP-00-23 Skewed Method Mix: a Measure of Quality in Family Planning Programs (Jane T. Bertrand, Janet Rice, Tara M. Sullivan & James Shelton)

WP-00-22 The Stymied Contraceptive Revolution in Guatemala (Roberto Santiso G. and Jane T. Bertrand)

WP-00-21 The Impact of Health Facilities on Child Health (Eric R. Jensen and John F. Stewart)

WP-00-20 Effort Indices for National Family Planning Programs, 1999 Cycle (John Ross and John Stover)

WP-00-19 Evaluating Malaria Interventions in Africa: A Review and Assessment of Recent Research (Thom Eisele, Kate Macintyre, Erin Eckert, John Beier, Gerard Killeen

WP-00-18: Monitoring the AIDS epidemic using HIV prevalence data among young women attending antenatal clinics: prospects and problems (Basia Zaba, Ties Boerma and Richard White)

WP-99-17: Framework for the Evaluation of National AIDS Programmes (Ties Boerma, Elizabeth Pisani, Bernhard Schwartländer, Thierry Mertens)

WP-99-16: National trends in AIDS knowledge and sexual behaviour in Zambia 1996-98 (Charles Banda, Shelah S. Bloom, Gloria Songolo, Samantha Mulendema, Amy E. Cunningham, J. Ties Boerma)

WP-99-15: The Determinants of Contraceptive Discontinuation in Northern India: A Multilevel Analysis of Calendar Data (Fengyu Zhang, Amy O. Tsui, C. M. Suchindran)

WP-99-14: Does Contraceptive Discontinuation Matter?: Quality of Care and Fertility Consequences (Ann Blanc, Siân Curtis, Trevor Croft)

WP-99-13: Socioeconomic Status and Class in Studies of Fertility and Health in Developing Countries (Kenneth A. Bollen, Jennifer L. Glanville, Guy Stecklov)

WP-99-12: Monitoring and Evaluation Indicators Reported by Cooperating Agencies in the Family Planning Services and Communication, Management and Training Divisions of the USAID Office of Population (Catherine Elkins)

WP-98-11: Household Health Expenditures in Morocco: Implications for Health Care Reform (David R. Hotchkiss, Zine Eddine el Idriss, Jilali Hazim, and Amparo Gordillo)

WP-98-10: Report of a Technical Meeting on the Use of Lot Quality Assurance Sampling (LQAS) in Polio Eradication Programs

WP-98-09: How Well Do Perceptions of Family Planning Service Quality Correspond to Objective Measures? Evidence from Tanzania (Ilene S. Speizer)

WP-98-08: Family Planning Program Effects on Contraceptive Use in Morocco, 1992-1995 (David R. Hotchkiss)

WP-98-07: Do Family Planning Service Providers in Tanzania Unnecessarily Restrict Access to Contraceptive Methods? (Ilene S. Speizer)

WP-98-06: Contraceptive Intentions and Subsequent Use: Family Planning Program Effects in Morocco (Robert J. Magnani)

WP-98-05: Estimating the Health Impact of Industry Infant Food Marketing Practices in the Philippines (John F. Stewart)

WP-98-03: Testing Indicators for Use in Monitoring Interventions to Improve Women's Nutritional Status (Linda Adair)

WP-98-02: Obstacles to Quality of Care in Family Planning and Reproductive Health Services in Tanzania (Lisa Richey)

WP-98-01: Family Planning, Maternal/Child Health, and Sexually-Transmitted Diseases in Tanzania: Multivariate Results using Data from the 1996 Demographic and Health Survey and Service Availability Survey (Jason Dietrich)

Rating Maternal and Neonatal Health Programs in Developing Countries

Submitted to the MEASURE Evaluation Project
University of North Carolina, Chapel Hill

by

The Futures Group International

August 2000

Subcontract No. 556125
HRN-A-00-97-00018-00

Rating Maternal and Neonatal Health Programs in Developing Countries

Abstract

Maternal and neonatal health services in 49 developing countries are compared, using national ratings provided by 10-25 experts per country on a variety of dimensions. These expert ratings cover both routine and emergency services at health facilities, access to these services for both rural and urban women, the likelihood that women receive particular forms of antenatal and delivery care, and such essential program elements as appropriate policy, health promotion, and training. Programs receive low ratings, but great variation exists across countries, especially with regard to access to services in rural areas. The ratings appear to confirm some aspects of the conventional wisdom about the quality of services. For instance, immunization services receive good ratings, while care for sexually transmitted infections receives abysmal ratings. However, some results are not quite as expected. Where comparisons are possible, the ratings are generally consistent with Demographic and Health survey data.

NOTE: This report was drafted by Rodolfo A. Bulatao and John A. Ross for the MNPI project at the Futures Group International. The work was suggested by the MEASURE Evaluation project, which provided major financial support; additional support was received from the Rockefeller Foundation.

Contents

Rating Maternal and Neonatal Health Programs in Developing Countries

Maternal and neonatal health services in 49 developing countries were rated in 1999 by experts on each country in order to provide a comparative assessment. This paper provides initial, descriptive results of this exercise.

Programs in developing countries are recognized to have numerous deficiencies, but some are certainly better than others. Exactly how services compare cross-nationally is difficult to say. Some outcome measures exist, such as maternal mortality rates, but cross-nationally comparable indicators of service adequacy and the adequacy of associated program elements are rare. The current exercise was conceived as an attempt to provide such measures, relying on expert judgments. A set of such measures might serve as a diagnostic tool, to identify strengths and weaknesses in a program; a planning tool, to indicate where and how program effort should be concentrated; and a research tool, to provide insights into contrasts between countries and relationships between types of services and health outcomes.

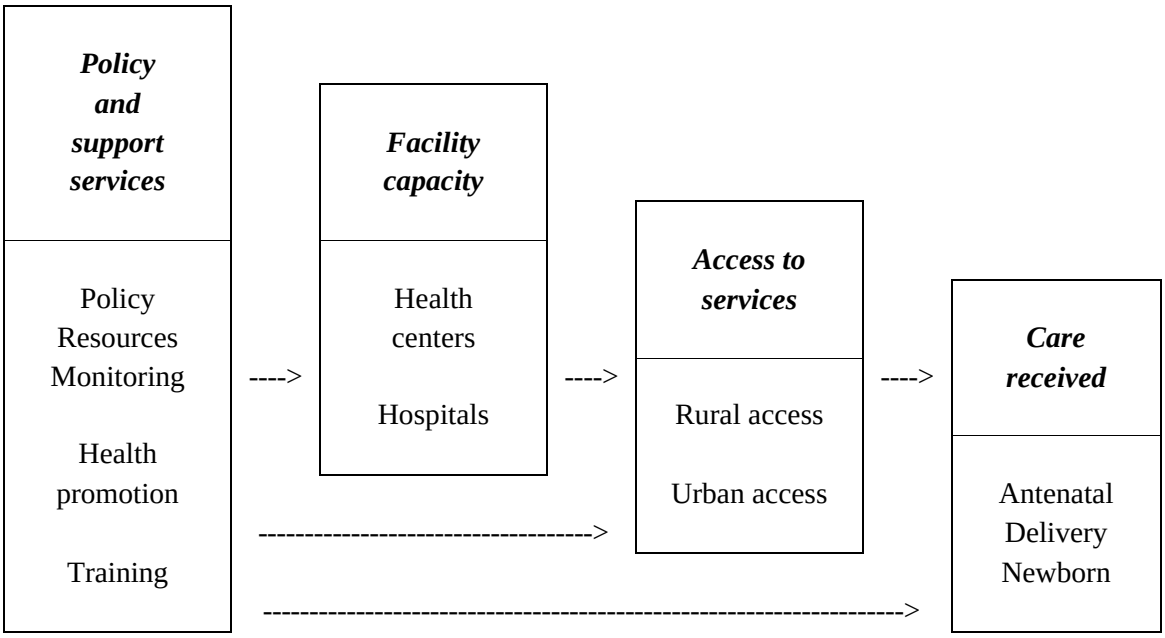
Various sets of indicators for maternal health and health services, and for reproductive health services more broadly, have been suggested, especially in the last decade (Maine 1991; Graham and Filippi 1994; World Health Organization 1994, 1997; Campbell et al. 1995; Prevention of Maternal Mortality Network 1995; Koblinsky et al. 1995; Bulatao and Shrestha 1996; National Research Council 1997; Maine et al. 1997; UNICEF et al. 1997). The indicators suggested in the literature were reviewed, and where feasible the authors were consulted in the process of designing the study.

Two important categories of maternal health services are those involving emergency care for such conditions as hemorrhage and obstructed labor and those involving routine and preventive care that can be provided, for instance, at antenatal visits. Recognizing that both types of care are essential, we cover both without attempting to decide their relative importance. We also cover neonatal care as a category of closely linked services. Some of the specific services we cover may not be the responsibility of the official maternal health unit in a health ministry but of other organizational units, but we still refer to all the services collectively as the maternal health program.

Country experts have been used before to rate programs in the area of family planning; a family planning effort scale has been applied cross-nationally five times: in 1972, 1982, 1985, 1994, and 1999 (Mauldin and Lapham 1985; Mauldin and Ross 1991; Ross and Mauldin 1996; Ross and Cooper-Arnold 2000). Program effort at the subnational level has also been investigated (Pham et al. 1999). The scale has been important in country comparisons (Mauldin and Berelson 1978) and in the analysis of the effects of such services (e.g., Bulatao 1986; Schultz 1994).

Expert ratings in the area of HIV/AIDS policy have been obtained for five Central American countries (Murgueytio et al. 1997; Nunez et al. 1998), and expert ratings for HIV/AIDS programs more broadly have been introduced recently (Stover et al. 2000). The use of experts to rate programs in the general area of reproductive health therefore has precedent, and the experience gained in these previous efforts was used to inform this exercise.

Figure 1. Aspects of maternal health programs evaluated in the study



Note: Family planning provision is another section of the questionnaire, along with the sections listed in the above blocks. It cuts across the blocks, by including postpartum and postabortion contraception, supplies on hand, and trained staff in place who can do IUD insertions (at health centers) or IUDs and sterilization (at district hospitals).

Method

The exercise was conducted by the Futures Group International, which developed a questionnaire, identified a consultant for each country to be studied, and coordinated her or his recruitment of experts to produce ratings collected in 1999 and early 2000. We will treat 1999 as the effective date for these ratings.

Measures

Each expert was asked to rate the national maternal and neonatal health program in one country on an 81-item questionnaire labeled the Maternal and Neonatal Program Effort Index (MNPI). The items refer to many related areas: essential obstetric services, antenatal care, newborn care, family planning, control of sexually transmitted infections, etc. Items were also included having to do with maternal and neonatal health policy and related health promotion, training, and research.

These items were grouped not by medical condition but with reference to the different stages involved in organizing and delivering care, from setting policy to attending to the patient. Figure 1 shows these stages and indicates, in broad terms, the expected relationships among them. Policy and various ancillary support services come first (though they appear last in the questionnaire). This stage includes policy development, resource provision, and monitoring and research--each a central function essential to

effective management and control of a program. Included at this stage are health promotion (also referred to as information-education-communication) and staff training. These ancillary program activities presumably enhance the capacity of health facilities. The most important types of these--health centers and district hospitals--are the ones covered in the questionnaire. Adequate facilities should contribute to sufficient access to services, which was rated separately for rural and urban areas. Finally, the last stage covers health care actually received, with distinctions made between care at antenatal visits, at delivery for the woman, and at delivery for the newborn.

Underrepresented in the diagram but also covered in the questionnaire is the provision of family planning, the one substantive area separately covered. Questions on family planning were not distinguished by stage. Though they mainly refer to the capacity of health centers and district hospitals to provide family planning, some of these questions also reflect degree of access. The diagram shows four main areas for the questionnaire and 11 or 12 subareas, depending on whether rural and urban access are treated together or separately. With the addition of family planning (at health centers and at hospitals), the main areas become five and the subareas either 13 or 14 (Appendix A).

As should be clear from Figure 1, the questionnaire focused on service inputs and such immediate outputs as the medical care that women actually receive. Collectively, these elements might be referred to as program effort. What the questionnaire does not and was not intended to cover are health outcomes.

The questionnaire was pretested in four countries (Ghana, Guatemala, Indonesia, and Zambia) and reduced to its final length. Since only a few of the questions were reworded, and then only in minor ways (see Appendix A), pretest data from these four countries are included in this report.

Experts rated services on a scale from 0 to 5, where 5 was meant to indicate that it was "completely true" that, for example, "All pregnant women have their labor monitored." The opposite end of the scale, a rating of 0, was meant to indicate that this was "completely false." In reporting results, these ratings are multiplied by 20, so that they run from 0 to 100, where 100 can be interpreted as "100 percent true." For one set of items, raters used a slightly different scale. In rating access to services, they were asked instead to indicate the percentage of pregnant women with "adequate" access to each service. These scores therefore also run from 0 to 100, though their interpretation is slightly different.

In addition to rating current programs, raters produced similar, retrospective ratings of programs as of three years previously (effectively in 1996). We focus mostly on current ratings but briefly summarize changes over three years.

Countries

The developing countries included in the ratings are listed in Table 1 by region. Of the 49 countries covered, 21 are in Sub-Saharan Africa (divided for our purposes into two regions), 13 in Latin America and the Caribbean, 10 in Asia (excluding the Middle East), and 5 in the Middle East and North Africa. The countries covered were limited by the available resources. Some coverage was desired of each developing region. At the same time, we intended to cover the largest developing countries as well as on countries of important policy interest.

Table 1. Developing countries studied by region,[a] percentage of regional population covered, and number of raters per country

Country	*Raters*	*Country*	*Raters*	*Country*	*Raters*
East and Southeast Asia		**South Asia**		**Middle East and North Africa**	
(pop. covered = 90%)		*(pop. covered = 98%)*		*(pop. covered = 46%)*	
Cambodia	22	Bangladesh	18	Algeria	10
China	10	India[b]	231	Egypt	10
Indonesia	16	Nepal	13	Iran	21
Myanmar	12	Pakistan	12	West Bank	15
Philippines	11			Yemen	15
Vietnam	24				
Latin America and the Caribbean		**Francophone Sub-Saharan Africa**		**Non-Francophone Sub-Saharan Africa**	
(pop. covered = 71%)		*(pop. covered = 58%)*		*(pop. covered = 93%)*	
Bolivia	25	Benin	24	Angola	17
Brazil	10	Congo	13	Ethiopia	17
Dominican Rep.	14	Congo, Dem. Rep. of	17	Ghana	19
Ecuador	25	Guinea	18	Kenya	16
El Salvador	12	Madagascar	10	Malawi	14
Guatemala	23	Mali	17	Mozambique	20
Haiti	19	Rwanda	25	Nigeria	17
Honduras	25	Senegal	14	South Africa	14
Jamaica	22			Sudan	17
Mexico	12			Tanzania	15
Nicaragua	22			Uganda	15
Paraguay	17			Zambia	21
Peru	16			Zimbabwe	13

[a] Not covered are the Transcaucasian and Central Asian republics of the former Soviet Union. Even with these and other exclusions, the proportion of the total developing-country population covered is 84 percent.
[b] For India only, ratings were obtained by state from different groups of raters. Total raters across states is given.

Though many developing countries were not covered, those that were include 84 percent of the population of the developing regions. Among those left out are the Transcaucasian and Central Asian republics of the former Soviet Union. The majority of the population of five out of six other developing regions is covered, the exception being the Middle East and North Africa, which is 46 percent covered.

Partly because of its size, India was treated differently from the other countries. State rather than national programs were rated, covering 14 of the 15 largest states (except Bihar) with about 85 percent of the national population. To obtain national ratings to compare with other countries, population-weighted averages of state ratings were calculated.

Table 2. Rater characteristics: Position, MD degree if any, and years of experience in maternal health

Position and MD degree, if any	*No. of raters*	*Mean years of experience*			
		National	*Provincial*	*District*	*Community*
All countries except India	**806**	**8**	**4**	**3**	**3**
Administrator MDs	121	9	4	3	3
Administrator non-MDs	58	10	5	3	3
Physicians	246	7	3	3	2
Nurses, midwives	92	10	4	2	2
Outside MDs	124	8	3	3	3
Outside non-MDs	163	9	3	2	3
Not reported	2	13	0	0	0
India	**231**	**3**	**8**	**6**	**6**
Administrator MDs	65	3	8	10	7
Administrator non-MDs	15	7	7	3	4
Physicians	28	0	7	5	4
Nurses, midwives	5	0	2	9	5
Outside MDs	68	2	9	4	6
Outside non-MDs	46	4	9	3	5
Not reported	4	0	6	10	5

Note: Years of experience are not necessarily additive across organizational levels. Outsiders are not regular employees of a health program but may advise it in various capacities.

Raters

Table 1 also shows the number of raters. There were at least 10 for each country and each Indian state, and as many as 25. These raters were deliberately diverse to represent different viewpoints. They included experts from ministries of health (who may be working in maternal and child health, hospitals, training, management information, or elsewhere), medical schools and universities, associations (of obstetricians and gynecologists, nurses, and midwives), nongovernmental and community organizations, and donors and related agencies (including resident staff of international organizations, other donors, and USAID cooperating agencies).

The characteristics of these raters with regard to positions held, MD degrees if any, and years of experience in maternal health are shown in Table 2. Raters were classified as administrators, service providers, and such outsiders to the program as university professors and consultants. Providers were 42 percent of the total across countries, excluding India where they were only 15 percent. Medical doctors, who could be practicing physicians, administrators, or outsiders, were 61 percent across countries but 71 percent for India.

The raters generally had many years of experience in maternal health. The average country rater had eight years of experience at the national level. The average Indian rater, on the other hand, also had eight years of experience at the provincial level (which was the level being rated for this country only). Raters appear to have at least as much experience working at subnational levels--in provinces, districts, and communities--but these numbers cannot easily be added up. An indeterminate number of raters reported working at more than one level at the same time. All in all, the raters recruited for this exercise appear to have been highly qualified, with substantial amounts of experience.

Analysis across raters indicates that more knowledgeable raters tended to skip fewer questions, and that raters holding different positions varied slightly, in comprehensible ways, in the level of their ratings (Appendix B). In what follows we generally consider means across all raters for a given country.

Program ratings

We consider, in order, ratings of:

- the capacity of facilities, specifically health centers and district hospitals, to provide maternal health services
- the degree of access to services in rural and urban areas
- the maternal and neonatal health care actually received
- family planning provision as a component of maternal health care
- policy and support services.

This corresponds to the order of stages in Figure 1, except that we discuss policy and support services last.

Facility capacity

Mean ratings across all countries of the capacity of health facilities are shown in Table 3. The items are also ranked, combining health center items and district hospital items, from the highest to the lowest mean. Standard deviations are shown to provide some idea of how greatly ratings vary across countries. (These do not incorporate the variance across raters.)

Ratings are not high. The scale for all items runs from 0 to 100, so that the midpoint of 50 should indicate that it is equally likely that a given item is true or false. That a health center has adequate antibiotic supplies (mean rating = 52) is therefore about equally likely to be true as to be false across all developing countries studied. What health centers are most likely to have the capacity to do is to administer antibiotics intravenously (rated 61). Unfortunately, this capacity is undermined by the fact that there is only a 52 percent chance that they will have the antibiotics on hand. Health centers tend not to use the partograph or have transportation arrangements for cases of obstructed labor and are especially unlikely to be able to offer manual vacuum aspiration or electric suction.

District hospitals generally score somewhat better than health centers, though they are far from perfect. Hospitals are best at doing the things that health centers are supposed to do (rated 67), but whether they have the capacity to provide blood transfusions (rated 52) is, on average across countries, close to even odds.

Table 3. Facility capacity to provide maternal health services: Mean ratings and item rankings, 49 developing countries

Having trained staff who can:	*Rank*	*Mean*	*S.d.*
At health centers			
Administer antibiotics intravenously	3	61	15
Manage postpartum hemorrhage	5	52	11
Have adequate antibiotic supplies	6	52	14
Manually remove retained placenta	7	49	16
Use partograph to determine when to refer	8	45	20
Have transport arranged for obstructed labor	9	43	15
Perform manual vacuum aspiration or electric suction	10	24	14
At district hospitals			
Provide all functions listed for health centers	1	67	10
Perform Cesarean section or other operative delivery	2	64	14
Perform blood transfusions	4	52	13

Variation across countries in facility capacity (as indicated by the standard deviation across countries per item) is largest with regard to the use of a partograph. This simple tool recommended by the World Health Organization (1993)--essentially a graph of cervical dilatation with "alert" and "action" lines--helps assess the progress of labor and determine when it is too prolonged and referral is needed. It has apparently been easily assimilated into health center procedures in some countries but is still largely unknown in others.

Though all these ratings of facility capacity appear low, are they still higher than they should be? This issue might also be raised with regard to other ratings that follow. Observers with international experience often have quite jaundiced views of developing-country services. Are they right? Without other cross-national data for comparison, it is difficult to tell. In the concluding section below we note two possible comparisons of these ratings with Demographic and Health survey data, neither of which indicates substantial discrepancies in rating level. But neither comparison involves facility capacity. In an appendix we also analyze rater variation, finding few systematic differences between raters and no particular reason to distrust any category of raters. But this still does not establish the validity of the scores they assign.

Some qualifications may, however, make these ratings appear more reasonable. First, the ratings represent averages across countries and, as we show below, variation between countries is substantial. Second, the ratings are based on local expert judgments. Local experts may be applying their own criteria, not necessarily as demanding as international standards. But they may also know more about how local programs actually work, and services can sometimes be provided in a manner that may strike international observers as disorganized and poorly managed but may still be effective. Third, the ratings apply only to existing facilities. Ratings can be high while facilities are few, serving a small proportion of those in need. The next set of ratings avoids this particular problem, measuring the access to services for all women.

Table 4. Access to maternal health services: Percent with access and item rankings, 49 developing countries

	In rural areas			*In urban areas*		
Pregnant women have adequate access to:	*Rank*	*Mean*	*S.d.*	*Rank*	*Mean*	*S.d.*
District hospitals open 24 hours	8	58	21	1	81	11
Antenatal care	9	56	18	2	80	11
Delivery care by trained professional attendant	11	44	18	3	75	15
Postpartum family planning services	12	36	21	7	61	18
Treatment for postpartum hemorrhage	13	35	17	5	69	16
Management of obstructed labor	14	33	17	4	69	16
Treatment of abortion complications	15	32	17	6	68	16
Provision of safe abortion services	16	21	18	10	45	22

Access to services

Access to maternal health services was rated separately for rural and urban areas, with raters asked to indicate what proportion of pregnant women had "adequate" access to each service (Table 4). Some access items are similar, though not identical, to those in the previous table.

Urban access is clearly much better than rural access, with raters estimating that 81 percent of urban women, on average, have access to a 24-hour district hospital. For rural women, the proportion drops to 58 percent. Responses on this item produce the smallest contrasts between countries (the smallest standard deviation) for urban areas but the greatest contrasts for rural areas.

Access to antenatal care is about as available as access to a 24-hour district hospital and shows a similar urban-rural contrast. Rural variation across countries again appears much greater than urban variation. For urban women, delivery care by a trained professional attendant is almost as accessible as antenatal care. For rural women, however, degree of access falls relative to antenatal care, being below 50 percent. For urban women, access to treatment for delivery complications (postpartum hemorrhage, obstructed labor, and complications of abortion) is available to a little over two out of three. For rural women, the corresponding proportion is half as high: one out of three.

Postpartum family planning is more accessible in urban than in rural areas. However, in comparison to other services, it receives somewhat more emphasis in rural areas. The least available service is safe abortion: fewer than half of urban women and only one in five rural women have access.

In some particulars, these access ratings agree with the preceding ratings of facility capacity. For instance, the greater urban access to district hospitals, and the better ratings of hospital capacity in the previous table, argue that urban women should have better access to specific forms of treatment, as appears to be the case. In addition, the marginally better access in rural areas to treatment for postpartum hemorrhage than for obstructed labor is consistent with the slightly better rating of capacity to manage postpartum hemorrhage than to transport for obstructed labor in the previous table. Overall, however, a close correspondence cannot be made. The wording of items differs, and comparisons are further

Table 5. Maternal and neonatal health care received: Mean ratings and item rankings, 49 developing countries

Care	*Rank*	*Mean*	*S.d.*
Pregnant women at antenatal visits:			
Receive needed tetanus injections	2	78	12
Are examined for hypertension and treated	7	70	11
Receive iron folate tablets for anemia	10	66	13
Are informed about danger signs	12	60	11
Are examined for syphilis and treated	17	51	20
Are offered voluntary HIV counseling and testing	19	30	14
Pregnant women at delivery:			
Are encouraged to immediately start breastfeeding	4	74	11
Are counseled on umbilical cord care	9	66	13
Are checked for hypertension, anemia, infection	11	60	11
Are seen by trained professional attendant	14	56	12
Can receive emergency obstetric care	15	55	12
Have labor monitored	16	52	14
Are scheduled for a checkup in 48 hours	18	41	15
Newborns:			
Are scheduled for subsequent immunizations	1	78	9
Receive DPT injection at 3 months	3	76	9
Have umbilical cord cut with clean blade	5	73	10
Are dried and kept warm	6	73	10
Have mouth and nasal passageways cleared	8	69	11
Receive prophylactic eye treatment	13	57	20

complicated by the geographical distribution of pregnant women relative to facilities, a factor about which we have no data.

Care received

The maternal and neonatal care actually received was rated in three areas: for pregnant women at antenatal visits, for pregnant women at delivery (whether in a health facility or not), and for newborns. Ratings on a scale of 0 to 100 are again considered, with a rating of 50 indicating even odds that appropriate care was or was not received (Table 5).

The ratings for care received are somewhat higher than the ratings of service capacity or access, indicating that pregnant women and newborns have somewhat better than even odds of receiving several types of care. This is partly because this set of items has proportionally less emphasis on obstetric emergencies and more emphasis on routine types of care.

The best odds for receiving care involve immunization. The three immunization items (scheduled immunizations, tetanus toxoid, and DPT) receive the highest ratings of all the items (76 to 78). These ratings still indicate that many are not covered, but their relatively high level must be a result of the vigor

Table 6. Family planning provision: Mean ratings and item rankings, 49 developing countries

	At health centers			*At hospitals*		
Have trained staff who:	*Rank*	*Mean*	*S.d.*	*Rank*	*Mean*	*S.d.*
Have contraceptive pills regularly in stock	3	66	14	2	67	15
Routinely offer family planning postpartum	6	59	14	5	62	13
Can insert intrauterine devices	7	57	19	1	71	12
Routinely offer family planning postabortion	9	51	14	8	56	14
Have progestin-only pills for breastfeeding women	10	49	21	--	--	--
Can offer sterilization to females	--	--	--	4	62	16
Can offer sterilization to males	--	--	--	11	36	22

of the worldwide immunization campaign.

At the opposite end, with very poor ratings, is voluntary HIV counseling and testing (rated 30). HIV is not considered a pressing problem in various developing countries covered, but where it is, ratings are still low. In addition, the other item having to do with sexually transmitted infections or STIs--examination and treatment for syphilis (rated 51)--is also close to the bottom of the ratings. Care for STIs appears to be a serious weakness, lagging well behind other areas.

In between immunization and STI care fall various other types of care. Newborn care appears somewhat better than antenatal or delivery care, with slightly less variation across countries. The weakest point regarding newborn care is prophylactic eye treatment, which also shows the greatest variation across countries, where newborn care is concerned.

As already noted regarding antenatal care, tetanus immunization is the most frequently available type of care and HIV and STI counseling the least. Hypertension receives slightly more attention at antenatal visits than iron folate supplementation, which in turn receives more attention than counseling about danger signs in pregnancy. This hierarchy of items is interesting, suggesting an edge for medical interventions over nutritional supplementation and simple counseling even at antenatal visits.

At delivery, the odds of having a trained professional attendant (rated 56) are slightly better than even. A pregnant woman has better odds that she will be encouraged to breastfeed, counseled on umbilical cord care, and checked for hypertension, anemia, and other infections--presumably by a less trained person. In an emergency, the odds for receiving care (rated 55) are again slightly better than even, but labor is less likely to be monitored (rated 52) to provide warning of an emergency. Monitoring, it might be noted, may be limited, given that capacity to use the partograph (see Table 3) has a lower rating. The lowest rated type of care is a scheduled checkup within 48 hours.

Family planning provision

Ratings of family planning provision combine elements of facility capacity, access, and care received (Table 6). These ratings range from 36 to 71.

District hospitals do better than health centers, and they do best at being able to insert IUDs (rated 71). They also tend to have contraceptive pills in stock, an area in which health centers are not too far

behind. Health centers do worst, among the items queried, at having progestin-only pills for breastfeeding women (rated 49). Hospitals do even worse, however, at providing male sterilization (rated 36).

Note that the likelihood that postpartum family planning is routinely offered is rated 56 for district hospitals and 51 for health centers. Table 4 showed that 61 percent of urban women have access to postpartum family planning services, suggesting that urban facilities that offer this service attend to proportionally more deliveries than other facilities. However, only 36 percent of rural women have such access, indicating that, whatever health centers and hospitals provide, substantial proportions of rural women do not have access to the facilities themselves, or at least to those facilities that are adequately staffed and equipped.

Policy and support services

Ancillary services are divided into these categories: policy, resources, monitoring and research, health promotion, and staff training (Table 7).

Broad policy is generally the strongest area. Having a basic policy (rated 72) and having a service director with a high rank in the bureaucracy (rated 67) both appear to be somewhat likely. Similarly, such other policy items as allowing appropriate personnel to provide services, developing policies through consultation with interested groups, and providing frequent public statements of support get better mean ratings than most other support-service items. The weakest areas (rated just above 50), where policy is concerned, are policies favoring treatment of abortion complications and active implementation of policies through high-level reviews and action plans.

The weakness of implementation is also reflected in poor scores in the area of resources. The odds are rated essentially even that the budget will be adequate (rated 48). In contrast, an active private sector is rated slightly more likely than not (rated 58).

To be effective, policy also requires active monitoring. In this area, item ratings vary. They are best for surveys of maternal events (rated 64), followed by statistical reporting systems (and their use for monitoring and decision-making) and then by facility listings. The lowest-rated item--unfortunately the one that could trigger the most immediate improvements in practices--is hospital reviews of all maternal deaths in the facility. In general, review and follow-up are particular weaknesses of developing-country maternal health systems. As earlier noted, the odds are barely even, or worse than even, for other reviews of services, whether high-level reviews or (the equivalent at the client level) scheduled client checkups. Providers may be too busy with clients (or too absorbed with competing activities) to reflect on their work in order to improve.

An important adjunct to the provision of services is educating the public about pregnancy complications, safe places to deliver, and harmful customs, but this receives relatively little attention. All the items on health promotion are closely clustered together below 50.

For staff training, on the other hand, the spread in item ratings is quite large. Hands-on training as part of medical curricula receives a relatively high rating, whereas training for new medical staff receives a low rating near the bottom among those listed in the table. Refresher training (within the last five years) has ratings intermediate between these. As one might expect, doctors are less likely to be subjected to either new-provider or refresher training than nurses and midwives, but the difference in each case is only 2 points.

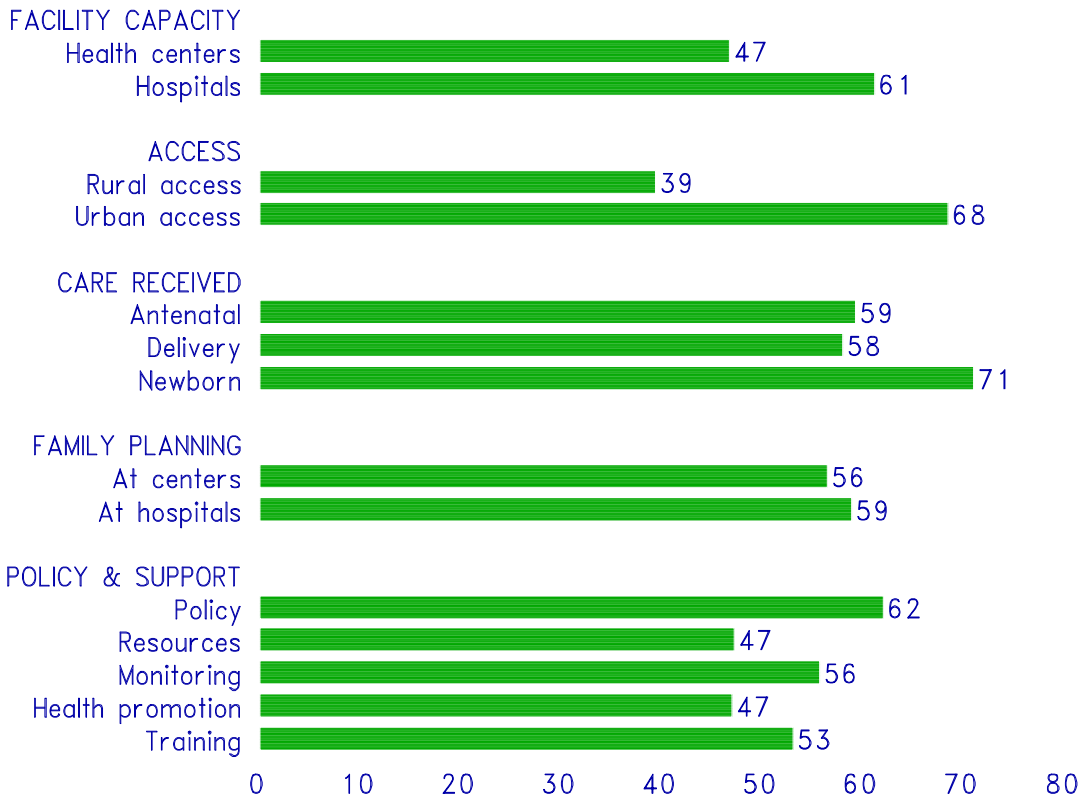
Item-group ratings

These results can be summarized using averages for each item group. As Figure 2 shows, the capacity of facilities to provide service is rated differently, on average, depending on whether health centers or district hospitals are being rated for services appropriate to each. A similar and even greater gap exists between access to services in rural areas and access in urban areas. On average across the developing countries, rural access is rated almost 30 points lower than urban access. Ratings across countries vary the most with regard to rural access.

Table 7. Policy and support services: Mean ratings and item rankings, 49 developing countries

Policy or support service	*Rank*	*Mean*	*S.d.*
Policy			
Adequate Ministry of Health policies	2	72	11
Service director at high administrative level	3	67	12
Appropriate personnel allowed to provide services	5	64	13
Policies developed through adequate consultation	6	64	12
High officials issue frequent statements of support	8	58	13
Policies favor treating abortion complications	12	55	15
High-level policy reviews and action plans	14	54	14
Resources			
Active private sector	7	58	14
Adequate budget	21	48	11
All services and drugs free	25	35	19
Monitoring, research			
Surveys provide data on maternal events	4	64	10
Statistical reporting system	9	57	14
Statistics used for decisions and strategy	10	56	13
Central monitoring and analysis of statistics	13	54	15
Updated listing of facilities	16	53	14
Each hospital reviews maternal deaths	17	50	16
Health promotion			
Ministry supplies educational materials	18	48	12
Community organizations educate public	19	48	10
Media-based education on complications	20	48	14
Media-based education on harmful practices	22	43	14
Training			
Medical curricula include hands-on training	1	73	12
Midwife-nurse refresher training within 5 years	11	55	13
Doctor refresher training within 5 years	15	53	14
New midwives and nurses trained in 6 months	23	43	16
New doctors trained for normal deliveries	24	41	15

Figure 2. Summary ratings for 49 developing countries



Health care actually received gets generally better ratings than either rural access or health center capacity, and on average at least as high ratings as hospital capacity. Newborn care has the highest ratings for any area at 71. As argued earlier, relatively higher ratings for care received may reflect the fact that these items emphasize emergency care less than do the facility capacity and access items.

Ratings on items relating to policy and support services are on average close to other ratings but vary across subcategories. Among these items, the tendency of developing countries to get better ratings for promises than performance--that is, for policy rather than for actually providing resources--appears clear. A tendency to pay less attention to health promotion than to other areas also appears in mean ratings below 50 in this area.

Contrasts and change

Countries

Countries vary a good deal in the ratings, and one way to show this is with the access ratings. Unlike other items, on which raters must make national judgments that combine possibly disparate urban and rural sectors, the access items allow them to make separate judgments. Combining these into national ratings is a matter of applying weights proportional to the population in each sector. As Table 8 shows, on this indicator countries vary greatly, for example Iran and Pakistan, in the same region, are almost at the opposite extremes. For convenience, countries in the table are grouped by level of national access, from "moderate" ratings of 70-89 down to "extremely weak" ratings of 10-29. Regions generally have a mix of

countries at different levels. The Francophone Sub-Saharan countries are the most tightly clustered, most of them receiving "very weak" national access ratings. Non-Francophone Sub-Saharan Africa, on the other hand, shows a broader range of ratings than most other regions, from South Africa at the top to Ethiopia at the bottom.

Regions

Despite the wide variation across countries, regions do not differ much when all items are averaged. The overall mean rating across all countries is 56. Four of the six regions have overall mean ratings within 1.5 points of this. South Asia has a mean rating 10 points lower, and East and Southeast Asia has a mean rating 5 points higher. Almost identical results would be obtained if means were first obtained for item groups (as in Figure 2) and these item-group means in turn averaged.

Combining items to give overall means may be misleading. As noted earlier, some items may reflect the adequacy of existing services regardless of how many women they serve. In comparing regions, therefore, we focus instead on means for item groups. These means for all countries combined were shown in Figure 2, and we examine now how regional means differ from these.

Figure 3 illustrates the variation from all-country item-group means for the two Sub-Saharan regions. These two regions contain 43 percent of the countries, and one might expect the means not to diverge too greatly from means for the entire sample. This is in fact what Figure 3 shows. Ratings for the non-Francophone countries as a group are all close to all-country means. Ratings for Francophone countries differ only slightly. The only areas where ratings diverge by 5 points or more, all for Francophone countries, are these: health center capacity, antenatal care, and policy, each rated marginally better than for all countries combined; and resources, family planning provision in hospitals, and rural and urban access, rated marginally worse.

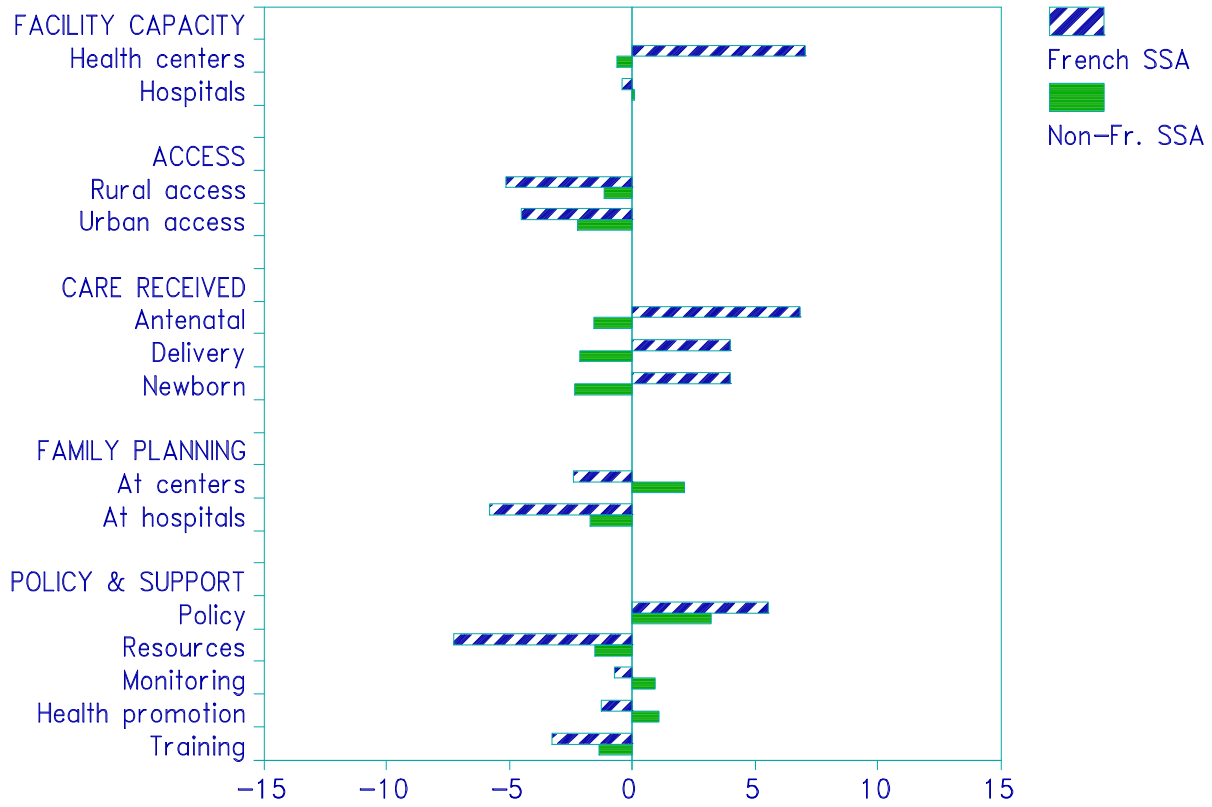
Figure 4 shows results for the two Asian regions. Here the contrasts are much greater. East and Southeast Asia has ratings above those for South Asia in every area. Though facilities in East Asia are rated just below those for all countries combined, East Asia has ratings better than the all-country average by 11-13 points on both urban and rural access and by 15 points on health promotion. Monitoring, training, and family planning in hospitals are also somewhat better than average in the region.

Table 8. National access ratings by country (combined rural and urban ratings weighted by population)

Level	*Latin America and the Caribbean*		*East and Southeast Asia*		*South Asia*	
Moderate (70-89)	Jamaica	83.1	China	75.4		
	Dom. Rep.	72.9	Vietnam	73.9		
	Peru	72.1				
Weak (50-69)	Mexico	66.1	Philippines	69.2	India	56.2
	Brazil	64.1	Myanmar	57.1		
	Paraguay	58.1	Indonesia	52.4		
	Ecuador	53.4				
	Nicaragua	50.6				
Very weak (30-49)	Honduras	49.7	Cambodia	33.0	Bangladesh	31.5
	El Salvador	47.9				
	Guatemala	40.4				
	Bolivia	39.1				
	Haiti	31.6				
Extremely weak (10-29)					Pakistan	24.6
					Nepal	16.9

	Middle East and North Africa		*Francophone Sub-Saharan Africa*		*Non-Francophone Sub-Saharan Africa*	
Moderate (70-89)	Iran	80.9			South Africa	73.3
	Egypt	74.5				
	West Bank	72.9				
Weak (50-69)	Algeria	66.4	Congo, Rep.	51.9	Zimbabwe	65.5
					Ghana	56.6
					Malawi	53.9
					Sudan	52.4
Very weak (30-49)			Benin	48.9	Tanzania	47.2
			Madagascar	48.1	Kenya	42.5
			Rwanda	44.3	Mozambique	42.2
			Mali	42.4	Nigeria	40.4
			Guinea	40.0	Uganda	40.3
			Senegal	39.7	Zambia	37.3
			Congo, DR	39.4	Angola	35.4
Extremely weak (10-29)	Yemen	29.4			Ethiopia	27.5

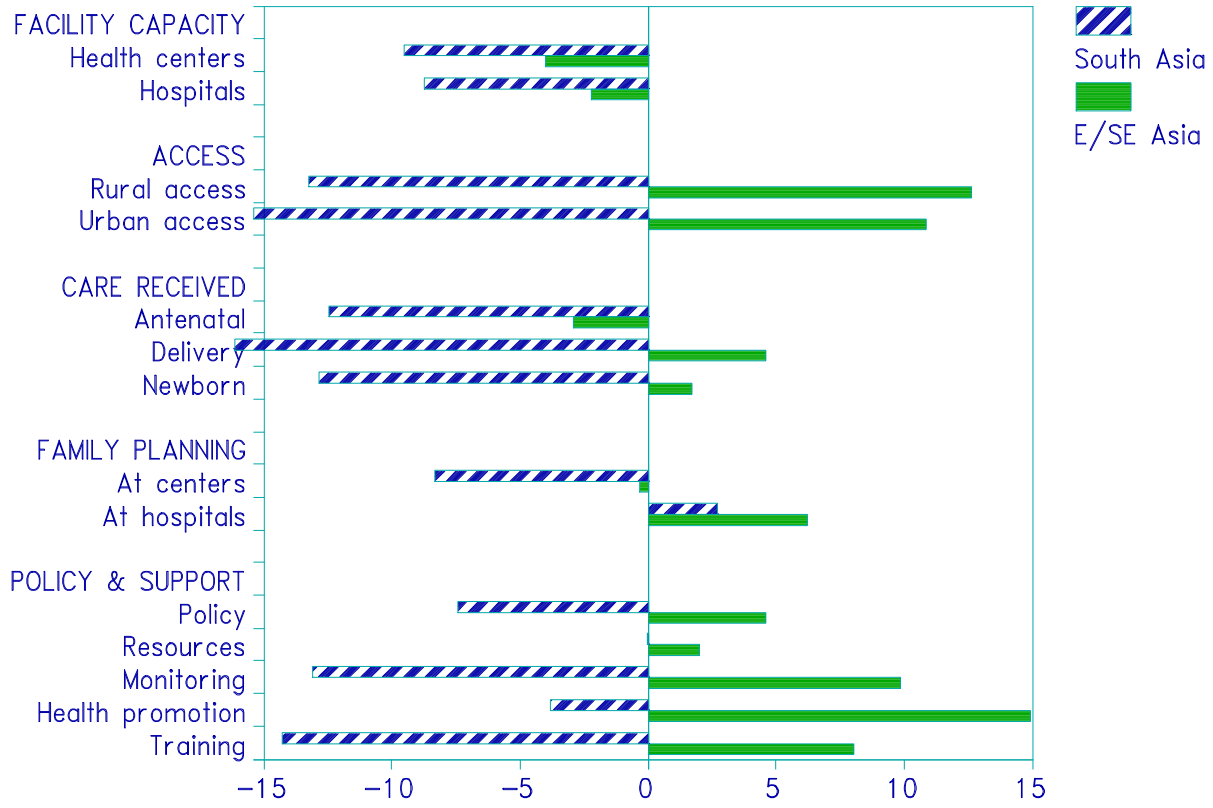
Figure 3. Deviations from mean developing-country ratings: Francophone and non-Francophone Sub-Saharan Africa



South Asia by contrast has ratings below the all-country means in all but one case. The exception is for family planning in hospitals, where South Asia is barely above the overall mean presumably because of provision of sterilization or IUDs. South Asia is 12-16 points below the overall mean on both rural and urban access, on care received in each of three areas, and on training and monitoring. In other areas, such as facility capacity, South Asia is not quite as far below the overall mean. Could underfunding be the cause of program weakness? The ratings suggest that South Asian programs are no weaker, in regard to resources, than programs in other regions. Nor can it be said that the private sector makes up for program weaknesses. On the one item explicitly referring to the private sector--having an active private sector that covers a substantial share of pregnancy and delivery cases--South Asia has lower mean ratings than any other region.

Figure 5 shows the last two regions, Latin America and the Caribbean and the Middle East and North Africa. Perhaps most surprising is the relatively moderate deviation of Latin American ratings from the all-country means. The Middle East and North Africa does do somewhat better than the all-country means at a few points. On resources, the Middle East and North Africa is 9 points better, but Latin America only 2 points better. The Middle East and North Africa does 12 points better on rural access and 7 points better on urban access, but Latin America does no better than the all-country means. However, on other indices--such as antenatal care received, training, monitoring, and health center capacity--the Middle East and North Africa does slightly worse than Latin America, so that across all items the mean ratings for these two regions are quite close.

Figure 4. Deviations from mean developing-country ratings: South Asia and East and Southeast Asia



As an alternative to overall mean ratings, on which the contrasts are somewhat unexpected, one might consider national access scores, earlier shown in Table 8 by country. Figure 6 shows these ratings by region, calculated in two ways: as simple averages of national access ratings and as averages weighted by the population in each country, the latter being referred to as regional access ratings. The two measures differ. Because larger countries tend to have better-rated services than smaller countries, regional access ratings are substantially higher than mean national access ratings, except in Sub-Saharan Africa. Thus, for all regions combined, mean national access is 51.3, while the weighted equivalent is 60.0.

Either measure contrasts with overall item means, showing much greater regional variation and a different ordering of regions. East and Southeast Asia and the Middle East and North Africa are first or second, and Latin America and the Caribbean is solidly in third. South Asia is still at the bottom on mean national access ratings, because Pakistan and Nepal are rated so low. But South Asia comes out ahead of Sub-Saharan Africa on regional access, because India, the largest country, is rated much higher than the others in the region.

Change

According to the raters, the weak scores for maternal health programs actually represent some improvement over past scores. Ratings as of three years previously indicate improvement over time, in all regions, on virtually all items.

Figure 5. Deviations from mean developing-country ratings: the Middle East and North Africa (MENA) and Latin America and the Caribbean (LAC)

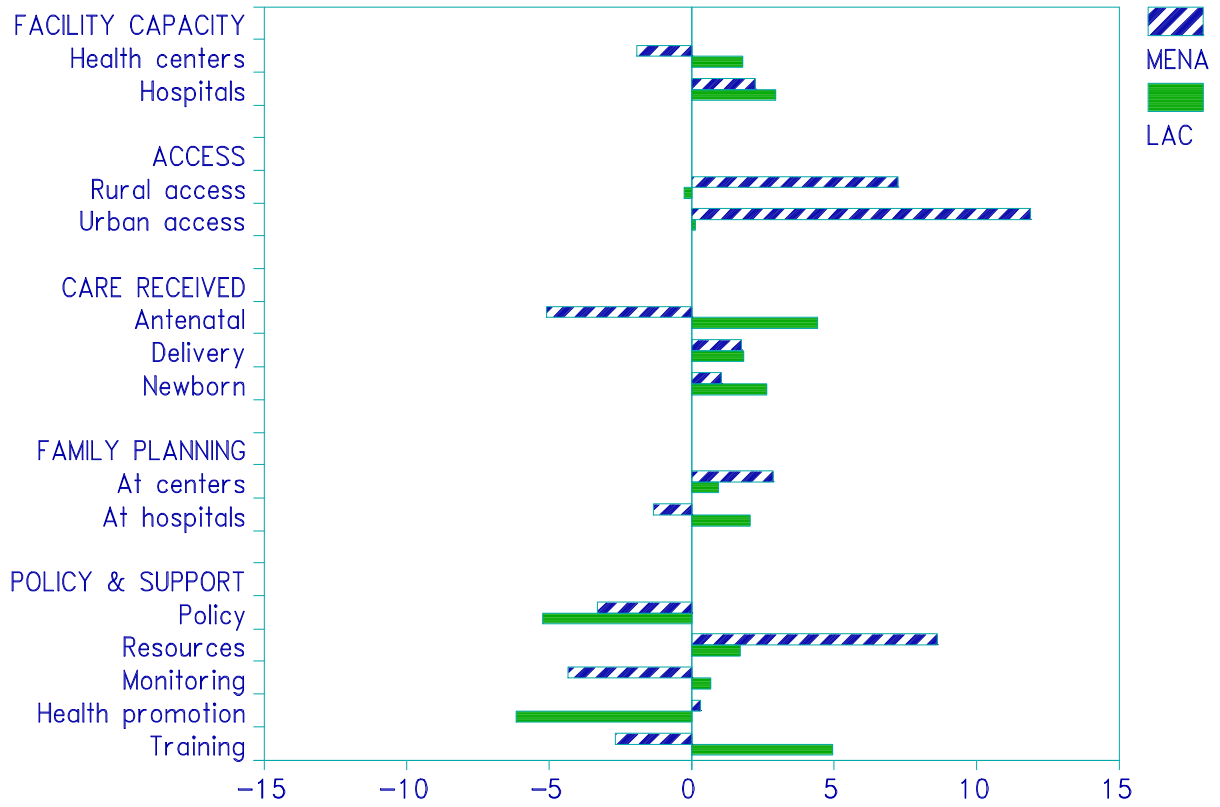
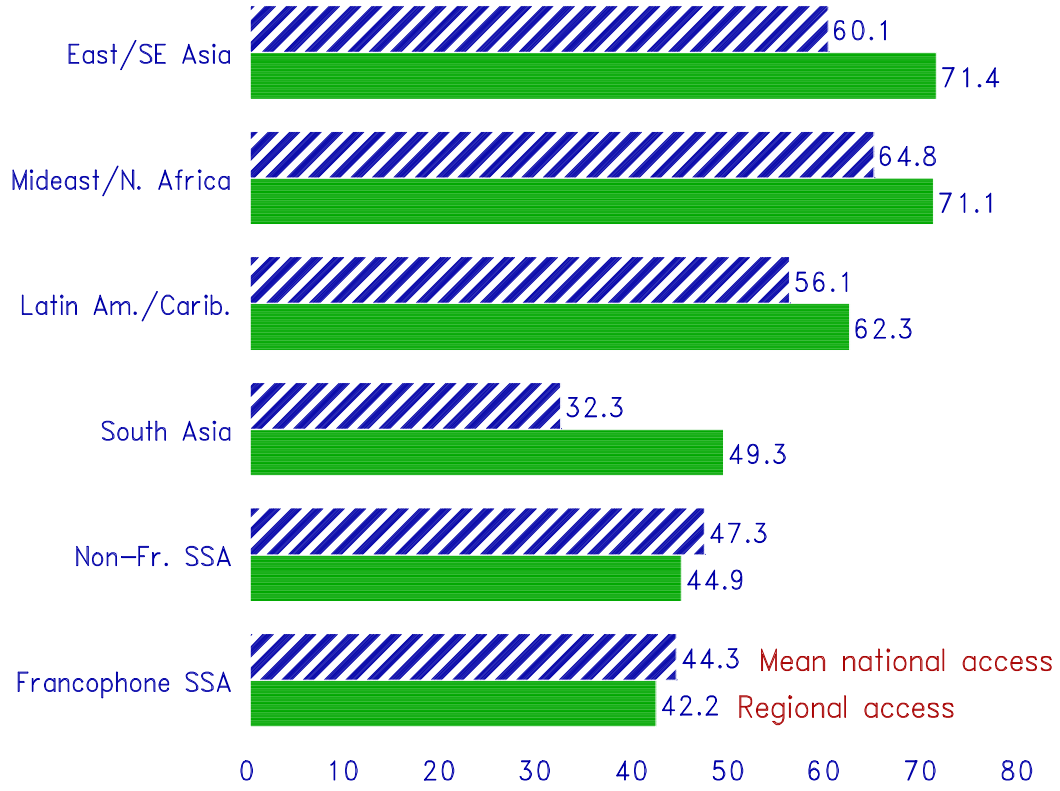


Table 9 shows the mean differences by region between 1996 and 1999 for each item group. A 10-point gain over three years is the norm. The level of gains might be somewhat overstated for methodological reasons. Any rater wishing to indicate some gain would have to make it a gain of at least 1 point on the scale from 0 to 5, or 20 points on the scale we are using here for descriptive purposes. Nevertheless relative gains for different countries or on different items should still be meaningful.

Gains have been greater overall in the two Asian regions than in other regions. The items on which East and Southeast Asia shows the most gains, absolutely and also in comparison to other regions, are those having to do with care received. The region also shows relatively good gains in policy and support services, especially health promotion. For South Asia, on the other hand, relative gains are better than average in policy and support services and also in hospital capacity.

In contrast to the Asian regions, Latin America and the Caribbean shows more limited gains than the developing-country average in policy and support services. On most other items, gains are roughly average for developing countries.

Figure 6. Regional ratings of access to services



The Middle East and North Africa posted relatively weak gains. Facility capacity did not improve as much as in other regions, and care received also lagged behind. The ratings indicate that the Middle East and Latin America were the overall leaders in 1996, barely ahead of East Asia. The services in both regions have improved, but not enough to avoid losing ground in comparison to other regions.

The two Sub-Saharan regions provide a contrast. Gains appear to have been slightly above average in Francophone Sub-Saharan Africa, slightly below average in non-Francophone countries. The difference between the two regions is consistent across item groups, with the Francophone gains always at least as high as and often a few points higher than the gains in non-Francophone countries. Across all items the difference in average gains is small: 11 points in Francophone countries, 9 points in non-Francophone countries. For some item groups, the difference is greater, particular for facility capacity and for care received. The Lusophone countries and Ethiopia are partly responsible for the limited gains in non-Francophone countries, but leaving them out would change the contrast between regions only minimally.

Across items rather than across regions, gains have been most limited, though still positive, with regard to rural and urban access: 7 and 8 points, respectively. Gains have been as limited with regard to resources. The resources area is the single area where individual item scores suggest regional deterioration: ratings for free services became marginally worse over the three years for both Sub-Saharan regions. The limited gains in access and resources suggest that other gains may have had less real impact than the numbers indicate.

Table 9. Mean reported change over the last three years, as estimated by raters, by region

Item group	*Combined*	*Latin America Carib.*	*East/ Southeast Asia*	*South Asia*	*Mideast/ North Africa*	*Franco-phone SSA*	*Non-Fr. SSA*
Facility capacity							
Health centers	12	12	14	13	8	15	11
Hospitals	12	12	13	18	9	16	10
Access							
Rural access	7	6	9	6	6	7	6
Urban access	8	6	11	8	5	9	8
Care received							
Antenatal	13	13	17	14	9	15	9
Delivery	13	12	17	14	10	14	11
Newborn	9	8	14	12	7	9	8
Family planning							
At centers	11	10	13	12	11	11	11
At hospitals	10	10	11	12	8	10	11
Policy and support							
Policy	12	9	15	17	12	15	11
Resources	7	6	10	12	9	7	6
Monitoring	11	9	14	15	7	12	11
Health promotion	12	8	17	17	12	12	10
Training	10	7	13	13	10	10	9

Conclusion

In summarizing these ratings of developing-country maternal and neonatal health programs, we compare the results with what we think is the conventional wisdom. We describe the conventional wisdom (CW) below in 24 propositions that we contrast with the actual findings. Our views of the conventional wisdom are not based on any hard data, and other observers may certainly hold different and more informed views. Nevertheless each proposition may serve as a useful foil against which to highlight the results of this inquiry.

The findings versus the conventional wisdom

CW-1. Maternal health programs in developing countries have serious deficiencies.

We made this statement at the outset, and the ratings generally confirm it. Raters estimate only a 56 percent likelihood, across countries, that the typical item in the menu of maternal health services will be adequate. That is unquestionably an unsatisfactory rating.

CW-2. Program adequacy varies among developing countries, though no country has fully satisfactory services.

As the contrast between Pakistan and Iran shows, developing-country programs receive widely varying ratings. The likelihood that a typical function is adequately handled ranges across countries from just over 30 percent to 75 percent. Even a 75 percent average indicates deficiencies, particularly since some ratings for individual items must be lower.

CW-3. By region, programs are weakest in Sub-Saharan Africa and in South Asia.

This is only partly true. South Asia indeed receives the lowest overall ratings. But overall ratings for Sub-Saharan Africa differ only slightly, on average, from those for Latin America or the Middle East. On the other hand, if one looks at regional access to services weighted by country populations, Sub-Saharan Africa is clearly behind the other regions. This seems to imply that the services that do exist in Sub-Saharan Africa are not necessarily worse than elsewhere, but given proportionally large rural populations, their coverage is more limited.

CW-4. Services tend to be considerably worse at health centers than at hospitals.

Raters considered separate items relating to maternal health services appropriate for health centers and hospitals. They gave health centers an overall rating of 47, as opposed to 61 for hospitals. Whether the items are properly calibrated to each level and represent a fair means of comparison is, however, difficult to say.

CW-5. Services are much worse in rural areas than in urban areas. Urban women have much readier access to needed services than rural women, especially in emergencies.

Service access ratings in rural areas average 39 percent (meaning 39 percent of rural women are estimated to have adequate access), as opposed to 68 percent for urban areas. Rural access ratings are among the most variable across countries, with mean ratings being as low as 13 percent in one case. Rural women are disadvantaged, relative to urban women, regarding access to antenatal care and to having a trained attendant at a delivery. They are even more disadvantaged, relative to urban women, regarding treatment of emergency obstetric conditions. Still, some countries have done considerably better than others in providing rural access, so the very low ratings in some countries might be remediable.

CW-6. Routine services for pregnant women are more likely to be provided than emergency services, especially in rural areas.

This seems generally the case, though there are exceptions. Items relating to care received, which mostly have to do with routine care, have higher ratings than items relating to facility capacity and access to services, which on balance focus more on emergency services. Those items on care received that deal with emergency obstetric services do receive low ratings, below other types of care. Nevertheless, there are some types of care that could be considered routine--such as STI and HIV counseling and scheduling a checkup--that receive even lower ratings than emergency care.

CW-7. Maternal health services are often provided on demand, as an apparent need arises, rather than individual pregnancies and individual women being systematically checked and followed up by health providers.

On the contrary, antenatal care in general receives relatively good ratings. Nevertheless, the one item that might reflect systematic follow-up, the scheduling of a checkup 48 hours after delivery, receives a mean score across countries of only 42. This may be symptomatic, though it does not prove the conventional wisdom.

CW-8. Routine newborn care is generally more adequate than care in obstetric emergencies.

Newborn care does receive the best overall scores of any set of items, averaging 71 across countries.

CW-9. Nutrition supplementation for pregnant women is underutilized.

The one item that might be said to relate to this, the use of iron folate tablets, receives a relatively good score (66) but still not as good a score as tetanus injections (78). It is arguable that more could be done within current health system capabilities.

CW-10. Breastfeeding has received more encouragement in recent years.

This appears clearly to be true. Encouragement of breastfeeding receives the best score (74) of all items relating to care received by pregnant women at delivery. Estimated improvement, over the last three years, is as high for this item (15 points) as for any other in the entire questionnaire.

CW-11. Immunization is a particular strength of these services.

The items relating to immunization do receive the highest ratings. Interestingly enough, these ratings appear to have improved less than those for the average item over three years, suggesting that the substantial improvements in immunization, for most cases, may have come before the mid-1990s.

CW-12. The provision of services relating to sexually transmitted infections receives low priority and is poorly organized.

The ratings support this. Two items relating to sexually transmitted infections and to HIV/AIDS are among those with the very lowest ratings.

CW-13. Due to substantial donor assistance for family planning, these services are usually more satisfactory than other maternal health services.

Actually family planning items receive ratings that are only around average, and some ratings are well below average. Rural access to postpartum family planning, for instance, is only 36 percent. The items on family planning were not selected to set standards more stringent than in other areas and probably do not do so. Perhaps family planning faces special obstacles, and the considerable assistance that has been available has on average only served to neutralize such obstacles. Or family planning may rely proportionally more on services provided outside health centers and hospitals, so that ratings based on services at these facilities overlook other areas of strength. Still another possibility is that the assistance provided for family planning has been of equal benefit in other areas of maternal health services, so that family planning services do not stand out.

CW-14. Male sterilization is considerably underutilized.

Certainly so, with a rating of only 36.

CW-15. A safe abortion is difficult to obtain in many developing countries.

This also seems to be clear. Rural access to safe abortion services is only 21 percent, and urban access is higher but still only 45 percent. Treatment for abortion complications is reported to be more accessible, but presumably would be less needed if abortion itself were accessible and safe.

CW-16. Official maternal health policies are becoming more constructive, but their effective implementation lags badly.

Adequate policy is one of the better areas, where various ancillary program activities are concerned. But in the policy area, an item having to do with high-level policy reviews and action plans receives the worst rating. Coupled with inadequate budgets (next point), this may signal weaknesses in implementation.

CW-17. Limited budgetary resources are a major obstacle to providing better services.

The likelihood of an adequate budget is close to 50 percent, not good but not that much worse either than other items. In addition, budgets judged to be more adequate in the Middle East and North Africa and less adequate in non-Francophone Sub-Saharan Africa do not seem to have produced any substantial gap in overall ratings. And resources no more limited than the average in South Asia go with much poorer than average access to services. Deficient current resources does not appear to be the single overriding cause of poorer services.

CW-18. Equipment and supplies, including transport for emergencies, are a special bottleneck and are typically a greater problem than having trained personnel. Maintenance of equipment and facilities is a notable weakness in the majority of countries.

There is not much evidence on this point. Antibiotic supplies and transport are weaknesses for health centers, but the gap is not excessive. There are no items relating to maintenance of facilities.

CW-19. Technical expertise from more advanced countries has upgraded many statistical systems for maternal health services. The actual use of statistics in planning is more uncertain.

One item relating to monitoring, having surveys of maternal events, gets a rating somewhat above average (64). But other items have only average ratings. If these have improved, they have not outstripped other areas, or substantial improvements have been confined to particular countries. The actual use of statistics in decisions is not rated substantially lower than the probability of having such statistics available.

CW-20. Health promotion and public education campaigns relating to maternal health receive insufficient attention.

All media-related items are rated, on average, below 50. Whether this is due to inattention, a lack of resources, or lack of skills to develop and implement decent health promotion plans, the weakness in this area appears evident.

CW-21. Academic training for maternal health service providers is often inadequate, but hands-

on in-service training is spreading.

In fact, medical curricula are reported to generally include hands-on training (rated 73). Refresher training receives an average rating, which has improved, over three years, at barely above the average pace of improvement.

CW-22. Doctors are harder to retrain than nurses and midwives.

Actually, ratings of training for doctors, whether refresher training or as new providers, are barely below those for training for nurses and midwives. What effect such training has, of course, has not been ascertained.

CW-23. Where public sector services are inadequate, private facilities provide more maternal health care.

There is no evidence of this across regions. If anything, the reverse appears to be the case. Where ratings are particularly low, in South Asia, an active private sector is also least evident.

CW-24. Maternal health care services have made little progress since the Cairo conference.

Over the last three years, raters estimate that adequacy has improved 10 points on the typical item. If their judgments are accurate, that would be fairly good performance, and, if sustained, could lead to substantial improvement over time. However, not all regions have been progressing at a similar rate. It remains to be seen why improvements have been smaller in some cases, and what the effect has been of such barriers as poor policy, limited donor support, economic problems, and such inherent social obstacles as deep social divisions.

Further questions

These conclusions rest on ratings from over a thousand experts, whom we relied on to provide an evaluation of national effort (or state effort in India) in maternal and neonatal health. How good are these ratings? A definitive answer is not yet available, but some indications exist that the ratings are at least reasonable.

Comparisons of raters indicate that those with more years of national experience in maternal health and those with MD degrees were more likely to complete the entire questionnaire rather than skipping items (Appendix B). This is what one would expect if raters drew on some actual knowledge of services. Raters who were program administrators and providers hardly differed in their ratings of different program areas. Raters outside a program, particularly MDs, appear to have given slightly lower ratings, though the differences were usually not significant. The level of rater agreement suggests no large biases in the ratings.

An additional way to look at the accuracy of the ratings is to compare two of the items on care received with other data. The Demographic and Health Surveys (DHS) asked national samples of women who attended to them at any birth in the five years before each survey and whether the women had received tetanus injections beforehand. Responses are available for 27 of the countries in our sample from surveys between 1994 and 1998.

The proportion of births with a trained attendant present, from DHS data, agrees well with the current ratings (as of 1999) for births attended. The correlation across countries is 0.70. The correlation is even stronger, at 0.83, with ratings as of three years previously--effectively for 1996. For the proportion receiving at least two tetanus injections, the correlations were also strong, at 0.62 for current ratings and 0.74 for ratings as of three years previously.

The level of ratings is also of interest. The mean percentage of births with a trained attendant across the 23 DHS countries is 55, virtually identical to the mean current rating for these countries of 56 and higher than the rating for three years previously of 43. From the DHS, the mean percentage receiving at least one tetanus injection is 69 percent and at least two injections 46 percent, below the current mean rating (for the same countries) of 77 for "needed" tetanus injections but, in the former case, almost the same as the mean rating for three years previously of 66. Many issues may be raised regarding how these comparisons should be interpreted and the reasons for agreement and disagreement, but they cannot be explored in detail here. What is important to note is simply that the ratings we have considered match, in some particulars, external data in these two areas.

More analysis is possible about how accurate the raters actually are. Additional data from other sources that might reflect system capacity might be considered. We have not investigated the relationship of program adequacy to morbidity and mortality outcomes. This report is meant to represent only an initial, mainly descriptive overview of a complex data set.

In describing the results we have at points used indices based on averaging items. The adequacy of this approach needs evaluation. Further analysis of the battery of items, beyond what we can consider in this report, might lead to improved indices. In addition, the differences between retrospective and current ratings deserve closer scrutiny.

Apart from more intensive analyses of the data, various questions are raised that require further investigation. Why do developing-country regions appear similar in some ways but so different in others? Is national access to services indeed the fairest way to make comparisons? Why are some countries rated much better than others, and do income, education, program leadership, or other factors account for the differences? Much remains to be learned from the way experts rate maternal and neonatal health programs in developing countries.

Acknowledgments

The basic data for this study came from 1,037 raters, who unfortunately cannot be acknowledged individually. The raters were recruited by consultant or consultant institutions in each country, who also made important contributions to the study. Advice on the design of the study and the questionnaire was received from a group that included Carla Abou-Zahr, Wendy Graham, Marge Koblinsky, Deborah Maine, Cindy Stanton, Patricia Stephenson, and Amy Ong Tsui. Assistance with data processing was provided by Katherine Abel and Katharine Cooper-Arnold.

Appendix A
Questionnaire Items and Variable Codes for the Maternal and Neonatal Program Effort Index (MNPI)

Maternal mortality and morbidity

I. All health centers have trained staff, in place, who can provide obstetric care:

1. CPPH Manage postpartum hemorrhage cases
2. CIVA Administer antibiotics intravenously
3. CPLA Perform manual removal of retained placenta
4. CMVA Perform vacuum aspiration of the uterus, using MVA (manual vacuum aspiration) or an electric suction device
5. CPAR Use a partograph to determine when to refer
6. CTRA Have transportation arrangements to quickly move a woman with obstructed labor to a district hospital
7. CSUP Have adequate antibiotic supplies on hand (sufficient supplies of the correct types)

II. All first referral facilities--district hospitals--have trained staff, in place, who can:

8. HCFU Provide all functions listed above for Health Centers
9. HBLO Perform blood transfusions (and have adequate supplies of safe blood on hand)
10. HCES Perform Cesarean section or other operative delivery (e.g. forceps delivery or craniotomy)

IIIa. What percentage of pregnant (rural) women have adequate access to:

11. RPPH Treatment for postpartum hemorrhage during or soon after delivery
12. ROBS Management of obstructed labor
13. RABC Treatment of abortion complications
14. RABO Provision of safe abortion services, or menstrual regulation
15. RANC Antenatal care during pregnancy
16. RDEL Delivery care by a trained professional attendant
17. RPFP Postpartum family planning services
18. RHRS District hospitals that are open 24 hours/day

IIIb. What percentage of pregnant (urban) women have adequate access to:

19. UPPH Treatment for postpartum hemorrhage during or soon after delivery
20. UOBS Management of obstructed labor
21. UABC Treatment of abortion complications
22. UABO Provision of safe abortion services, or menstrual regulation
23. UANC Antenatal care during pregnancy
24. UDEL Delivery care by a trained professional attendant
25. UPPH Postpartum family planning services
26. UHRS District hospitals that are open 24 hours/day

Maternal health

IV. At antenatal visits, all pregnant women:

27. VIRO Receive iron folate tablets for anemia
28. VHYP Are both examined for hypertension, and treated as needed
29. VSYP Are both examined for syphilis, and treated as needed
30. VTET Receive needed tetanus injection(s)
31. VSIG Are informed about danger signs of obstetric and newborn complications and are assisted in planning for any emergency
32. VHIV Are offered voluntary counseling and testing for HIV

Delivery and neonatal care

V. At delivery, all pregnant women:

33. DATT Are seen by a professionally trained attendant (either at home or in a facility)
34. DMON Have their labor monitored
35. DHYP Are checked for signs of hypertension, anemia, or infection
36. DEOC Are able to receive emergency obstetric care as needed
37. DCKU Are provided an appointment for a check-up within 48 hours of delivery
38. DBFE Are encouraged to immediately start breastfeeding their newborn
39. DUMB Are counseled on umbilical cord care

VI. For newborn care, all infants whether delivered at home or in a facility:

40. NMOU Have their mouth and nasal passageways cleared
41. NDRY Are dried and kept warm immediately after birth
42. NEYE Receive prophylactic treatment for their eyes.
43. NUMB Have their umbilical cord cut with a clean blade
44. NDPT Receive a DPT injection at 3 months
45. NIMM Are scheduled for subsequent immunizations

Provision of family planning

VII. All health centers:

46. FABO Routinely offer family planning after abortion cases
47. FPPT Routinely offer family planning at postpartum visits
48. FPIL Have contraceptive pill supplies regularly in stock
49. FPRG Have progestin-only pill supplies for breast-feeding women
50. FIUD Have trained staff, in place, who can insert intra-uterine devices

VIII. All first referral facilities--district hospitals:

51. GABO Routinely offer family planning after abortion cases
52. GPPT Routinely offer family planning at postpartum visits

53. GPIL Have contraceptive pill supplies regularly in stock
54. GIUD Have trained staff, in place, who can insert intra-uterine devices
55. GFST Can offer sterilization to female clients
56. GMST Can offer sterilization to male clients

General supporting functions

IX. Policies toward safe pregnancy and delivery

57. PMOH Ministry of Health policies toward pregnancy and delivery services are adequate
58. PCNS Policies are developed through adequate consultation with interested parties such as other ministries, NGOs, private practitioners, women's groups
59. PPER Policies are reasonable and fair concerning which personnel can provide maternal health services (e.g. trained midwives can perform a wide range of medical procedures)
60. PABC A favorable policy exists toward the treatment of complications of abortions, including complications seen from illegal abortions
61. PREV Policies are vigorously implemented through regular high-level reviews and updated action plans
62. PLVL The director of services for maternal health is placed at a high administrative level
63. PSOS High officials in the government, including the Ministry of Health, issue frequent statements to the press and public to support improvements for safe pregnancy and delivery

X. Resources

64. EBUD The government budget for safe pregnancy, delivery, and postpartum care (for facilities, personnel, supplies, etc.) is adequate for the needs, whether from the Ministry of Health, provincial government or donor support
65. EFRE All services and drugs are provided free to all clients
66. EPRI The private sector (doctors, midwives, clinics, maternity homes) is active and covers a substantial share of pregnancy and delivery cases

XI. Information, education, communication

67. IEDU The national program uses the mass media to educate the public about symptoms of pregnancy complications and safe places for childbirth
68. IHRM Also, the national program uses the mass media to educate the public about harmful home practices for pregnancy care, delivery, and postpartum care (home remedies and birthing customs, etc.)
69. IORG Community-level organizations take part in systematic programs to educate the public about safe pregnancy and delivery
70. IMAT The Ministry of Health supplies adequate educational materials (posters, pamphlets, etc.) to delivery facilities to instruct clients about safe practices

XII. Training arrangements

71. TCUR Medical curricula include hands-on clinical training in obstetric care, including management of several deliveries
72. TRMI All midwives and nurses in health centers have received refresher training for safe

pregnancy and delivery care within the last 5 years

73. TRMD Doctors in health centers have received refresher training for safe pregnancy and delivery care within the last 5 years
74. TNMI Newly hired midwives and nurses for health centers receive training for safe pregnancy and delivery care within the first 6 months
75. TNMD Newly hired doctors receive special in-service training for normal deliveries

XIII. Monitoring, evaluation, research

76. MSTA A routine statistical system (using facility-based information) provides good periodic information on supplies, personnel, deliveries, Cesarean sections, and cases of complications
77. MMON Staff at the national level regularly monitor and analyze results from the routine statistics (above)
78. MSUR Recent surveys provide data on maternal events (pregnancies, deliveries, attendants and sites for deliveries, estimates of maternal deaths, etc.)
79. MFAC An updated listing exists of facilities that can treat obstetric emergencies
80. MDEC Ministry administrators systematically use statistical information for decisions and reconsideration of strategies for reducing maternal mortality
81. MREV Each hospital follows a regular procedure to review and learn from every case of a maternal death in the facility

Variations in the questionnaire

The questionnaire for four pretest countries had 19 additional questions, which were later dropped. The other questions were worded as above, except for five that used wording that was later revised. The original wording:

42. NEYE Have their eyes checked for infection
57. PMOH Policies toward pregnancy, delivery, and aftercare are favorable and constructive
68. IHRM The media are used to educate the public specifically about harmful home practices for pregnancy care, delivery, and aftercare (home remedies and birthing customs, etc.)
71. TCUR Medical curricula include training in obstetric care for all health personnel who will attend deliveries.
79. MFAC An updated listing exist for facilities that can handle obstetric complications

Appendix B
How Raters Differ

Do different raters assess maternal and neonatal health services differently? The answer has implications for the validity of ratings and for whose ratings, if any, one should trust. The data on raters are limited, but we consider all that are available.

Each rater reported his or her official position, an MD degree if any, and years of experience. Official position might make a difference because those with authority over a program might have a reason to overstate its accomplishments. Those outside a program, on the other hand, could be more objective or could be less informed. An MD degree could be taken to indicate a more precise understanding of maternal and neonatal health service requirements. Years of experience might affect ratings in various ways. For instance, those with more experience might become disillusioned and more critical, or they might become more invested in a program and more likely to approve of it. These are only speculations, however, and need to be assessed against the data.

Method

Raters reported their number of years of experience in maternal health services at national, provincial, district, and community levels. National years of experience will be used in analysis of those rating countries, provincial years of experience for those rating Indian states. We divide the raters into groups of roughly equivalent size: those reporting 0 years of experience (these raters always have some experience at some other level), 1-5 years, 6-15 years, and 16 or more years. Total years of experience at any level of organization might be of interest but was not ascertained. In an indeterminate number of cases, reported experience at different levels overlaps and is not additive.

Raters are sorted by position into six categories: administrators with MDs, other administrators, physicians working as providers, nurses and midwives, outside experts with MDs, and outside experts without MDs. With 806 country raters, each of these six groups has at least 50 members. For the raters of Indian states, however, the nurse-midwife group has only 5 members, and the total is only 231.

Besides mean ratings across these rater categories, we will report analyses of covariance, with controls to reflect the fact that raters were assessing different programs. We will also look at variation across raters on specific items.

Items skipped

Raters were not required to fill in every item but were advised that, if they had insufficient knowledge, they could leave items blank. Half of the raters left at least one item blank, and some skipped a substantial number (Table B1). One reason for skipped items was an error in the questionnaire for Tanzania, which reached the raters minus the last page. This is part of the explanation why country raters were more likely to have skipped more than one or two items, in contrast to raters for Indian states, who, when they skipped items, generally skipped only one or two.

Among country raters, those who reported no national experience were somewhat more likely to skip at least one item. Those with more than 15 years of experience (as well as those with 1-5 years) were

Table B1. Percentage of raters skipping questionnaire items, by years of experience and position

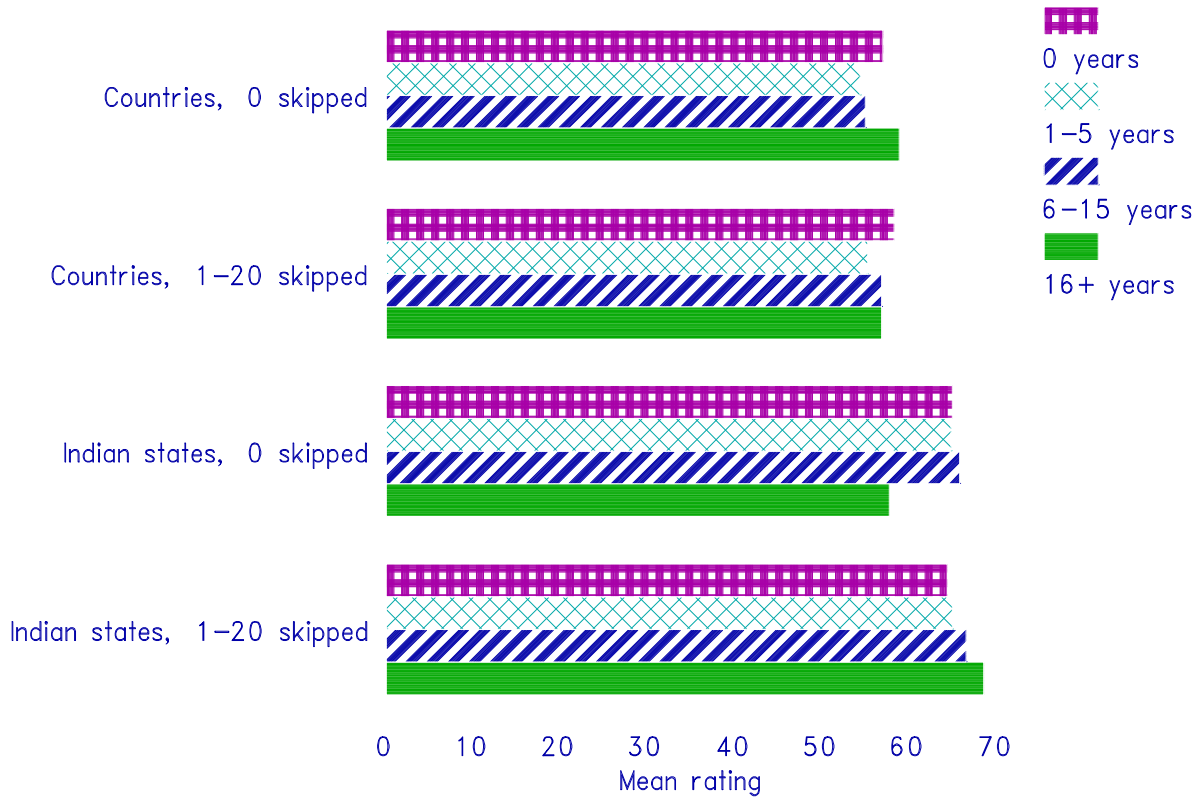
	Countries					*Indian states*				
	No. of items skipped					*No. of items skipped*				
Rater category	*0*	*1-2*	*3-10*	*11+*	*(N)*	*0*	*1-2*	*3-10*	*11+*	*(N)*
All	51	15	18	16	(806)	51	34	11	4	(231)
Years of experience										
0	44	17	21	18	(209)	51	34	13	1	(70)
1-5	57	16	12	14	(180)	55	31	10	4	(49)
6-15	51	14	19	15	(249)	51	34	11	3	(61)
16+	56	14	18	11	(140)	51	41	5	2	(41)
No data	43	14	14	29	(28)	20	20	30	30	(10)
Position										
Administrator MD	55	17	18	11	(121)	48	37	9	6	(65)
Administrator	36	21	21	22	(58)	67	27	0	7	(15)
Physician	57	14	14	15	(246)	36	43	18	4	(28)
Nurse, midwife	41	23	18	17	(92)	40	40	20	0	(5)
Outsider MD	56	11	20	13	(124)	60	31	7	1	(68)
Outsider	49	13	20	18	(163)	48	30	17	4	(46)
No data	100	0	0	0	(2)	25	50	25	0	(4)

slightly more likely than others to complete the entire questionnaire. Such tendencies as these were less evident among raters for Indian states, though those with more than 15 years of experience were clearly less likely to skip more than two items. (Taking national experience into account for Indian raters would make the picture slightly more consistent.) In addition, among both country raters and raters for Indian states, those who did not report their years of experience skipped items considerably more frequently.

Among country raters, those with MDs were either more thorough, more knowledgeable, or more opinionated: they were more likely than those without MDs to have turned in complete questionnaires. Among raters for Indian states, this contrast did not hold. Instead, complete questionnaires were most likely from administrators without MDs, followed by outside MDs. Practicing physicians among the Indian raters were the least likely to complete all the items.

Some of the contrasts among country raters are in line with what one might expect or hope: that those with more experience answer questions more fully, and that MDs in particular provide more information than others. The contrasts among Indian raters, however, are less expected and less interpretable, possibly because, with smaller numbers in each group, figures are less stable.

Figure B1. Mean ratings by years experience and number of skipped items



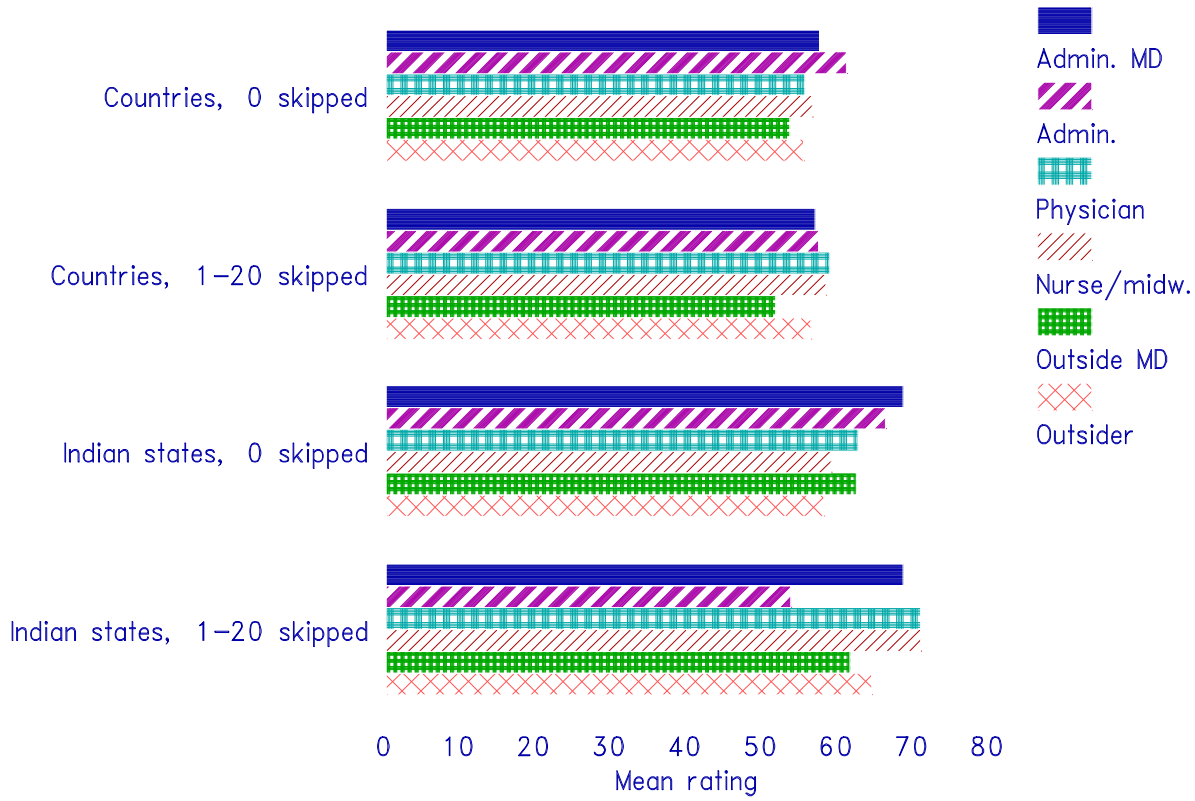
Overall rater variation

We consider now the actual ratings, beginning with a simple average across all items. If we want to deal with more than half of the raters, however, we need to fashion a rule for what to do when items have been skipped. We will divide the raters into two groups: those who filled out all questions and those who skipped up to 20 of the 81 questions on current services. This leaves out 29 raters who skipped even more questions. For those who skipped between 1 and 20 questions, the mean score is calculated across the items they did respond to.

Figure B1 shows how mean ratings across all items (or all answered items) vary, for raters of countries and of Indian states separately, depending on years of experience and whether any items were skipped. For country raters, variations appear to be small. For raters of Indian states, slightly lower ratings appear for those with 16 or more years of experience, as long as they filled out the questionnaire entirely. If they did not, then their ratings appear slightly higher. Aside from this effect, having skipped items does not affect mean ratings.

Figure B2 shows similar mean ratings by rater position. Here the variations may be more meaningful. Outside MDs give lower country ratings, whether or not they filled out the questionnaire entirely. Among raters for Indian states, mean ratings vary more but not in any immediately interpretable pattern. Administrators with MDs appear to give the highest ratings overall, but aside from this, the contrasts are not clear.

Figure B2. Mean ratings by rater position and number of skipped items



These comparisons do not take into account the specific countries (or states) being rated. Because the mix of raters may vary country by country, comparisons within countries are desirable. However, with only 10-25 raters per country, such comparisons would involve quite small samples. Instead, we will run analyses of covariance of mean ratings, using the categories of raters shown in the figures, and adding as a covariate the mean rating per country across raters. The covariate effectively controls for the fact that raters are considering different countries. We will also look at regional variation. Average regional ratings should not differ once country means are controlled. However, differences between raters in different categories may still be greater or smaller across regions, just as differences between raters contrast between the country comparisons and in the Indian state comparisons.

Table B2 shows the results. Mean country rating, as one would expect, is a significant covariate. But after it is controlled for, items skipped and years of experience do not come close to having a significant effect. Rater position is not significant either in country ratings but is significant in ratings of Indian states. As expected, region has no first-order effect, given that country means have been controlled. What is of interest is that none of the interactions involving region are significant either. This means that, for example, any gap between ratings of outside MDs and other raters is consistent, or at least does not vary significantly, across regions.

Table B2. Analyses of covariance of mean ratings, for countries and for Indian states

Source of variation	*Sum of squares*	*df*	*Mean square*	*F*	*p<*
Countries					
Covariate: Country mean	55,327	1	55,327	379.64	0.000
Main effects	2,154	14	154	1.06	0.396
Skipped items	219	1	219	1.51	0.220
Years experience	371	3	124	0.85	0.467
Position	1,253	5	251	1.72	0.128
Region	50	5	10	0.07	0.997
Two-way interactions	7,642	68	112	0.77	0.910
Skipped x Years	61	3	20	0.14	0.936
Skipped x Position	135	5	27	0.19	0.968
Skipped x Region	758	5	152	1.04	0.393
Years x Position	2,803	15	187	1.28	0.207
Years x Region	1,484	15	99	0.68	0.807
Position x Region	2,918	25	117	0.80	0.743
Explained	65,123	83	785	5.38	0.000
Residual	95,020	652	146		
Total	160,143	735	218		
Indian states					
Covariate: Country mean	8,976	1	8,976	61.15	0.000
Main effects	2,329	9	259	1.76	0.077
Skipped items	27	1	27	0.19	0.668
Years experience	145	3	48	0.33	0.804
Position	2,016	5	403	2.75	0.020
Explained	11,306	10	1,131	7.70	0.000
Residual	30,238	206	147		
Total	41,543	216	192		

Deviations from the grand mean, before and after adjusting for other effects, are shown in Table B3. Given the lack of significant effects in country ratings, we expect the adjusted deviations to be small, and they are. The greatest adjusted deviation is due to lower ratings by outside MDs, which may be worth further attention. The significant effect due to position in the Indian state ratings is shown to be somewhat complex, related especially to higher ratings given by administrators with MDs and by physicians and lower ratings given by outsiders without MDs.

Table B3. Deviations from grand mean ratings due to main effects, unadjusted and adjusted for covariate and other main effects

	Countries			*Indian states*		
Main effect	*Unad-justed*	*Ad-justed*	*N*	*Unad-justed*	*Ad-justed*	*N*
(Grand mean)	(56.4)			(64.7)		
Skipped items						
None	-0.5	-0.5	401	-0.9	-0.3	114
1 to 20	0.6	0.6	335	1.0	0.4	103
Years experience						
0	1.2	-0.2	195	-0.5	-0.8	69
1-5	-1.5	-0.8	173	0.1	0.2	49
6-15	-0.7	-0.1	233	1.7	1.3	58
16+	1.5	1.5	135	-1.8	-0.6	41
Position						
Administrator MD	0.8	1.5	111	3.7	3.8	62
Administrator	2.2	1.3	50	-2.0	-2.5	14
Physician	0.5	-0.4	234	3.5	3.0	27
Nurse, midwife	1.3	1.5	89	1.5	0.3	5
Outside MD	-3.2	-2.5	111	-2.3	-1.7	64
Outsider	-0.4	0.1	141	-3.4	-3.9	45
Region						
Latin America/Caribbean	0.4	0.4	226	--	--	--
East/Southeast Asia	4.7	-0.2	93	--	--	--
South Asia	-15.0	-0.4	42	--	--	--
Middle East/North Africa	1.9	-0.3	70	--	--	--
Francophone Sub-Saharan Africa	0.8	0.0	132	--	--	--
Non-Fr. Sub-Saharan Africa	-0.8	-0.1	173	--	--	--

-- Not applicable.

Rater variation by item group

The contrasts among raters by position can be explored further by looking at variation in ratings not for all items together but for item groups. Table B4 summarizes separate analyses of covariance for these ratings, again using the country means (or state means for India) as a covariate. We include only rater position as a main effect. (Adding the other, insignificant main effects in Table B2 would not change the results.)

For ratings of countries, rater position has an effect significant at 0.05 for 2 out of 14 item groups and is close to this level of significance for 3 other groups. For ratings of Indian states, rater position is significant for exactly half the item groups and close to being significant for 3 other groups.

To interpret these effects, we look at adjusted deviations from country means in Table B5. One pattern is consistent: the ratings of outside MDs are always below the average rating for an item group.

Table B4. Analyses of covariance of item group ratings: F-statistics for effect of rater position

	Countries		*Indian states*	
Item group	*F*	*p<*	*F*	*p<*
Health center capacity	2.16	0.057	2.76	0.020
Hospital capacity	0.51	0.766	0.21	0.958
Rural access	0.62	0.686	1.69	0.141
Urban access	4.45	0.001	2.76	0.019
Antenatal care	2.00	0.076	2.57	0.028
Delivery care	2.68	0.020	1.59	0.164
Newborn care	2.04	0.071	2.18	0.057
Fam. plan. at centers	1.58	0.165	3.72	0.003
Fam. plan. at hospitals	0.76	0.578	2.37	0.040
Policy	0.82	0.537	3.91	0.002
Resources	0.91	0.473	2.07	0.071
Monitoring	1.25	0.285	3.31	0.007
Health promotion	1.81	0.109	2.18	0.057
Training	1.44	0.210	1.12	0.353

Note: The covariate was the mean country rating (or state rating for India) for the given group of items.

The difference is not large--generally between 2 and 4 points below the average--but it is unmistakable. Outside MDs are most pessimistic, relative to other raters, on health center capacity and monitoring, and also rather pessimistic on delivery care and newborn care. Other difference between raters are less consistent across items, involving individual item groups and particular rater groups. Administrators without MDs are particularly positive about health promotion and training, especially in contrast to outside MDs and nurses and midwives. However, these two groups agree in giving antenatal and delivery care slightly higher ratings than other raters. Administrators without MDs also rate urban access more negatively, in contrast especially to administrators with MDs, who are considerably positive on these items.

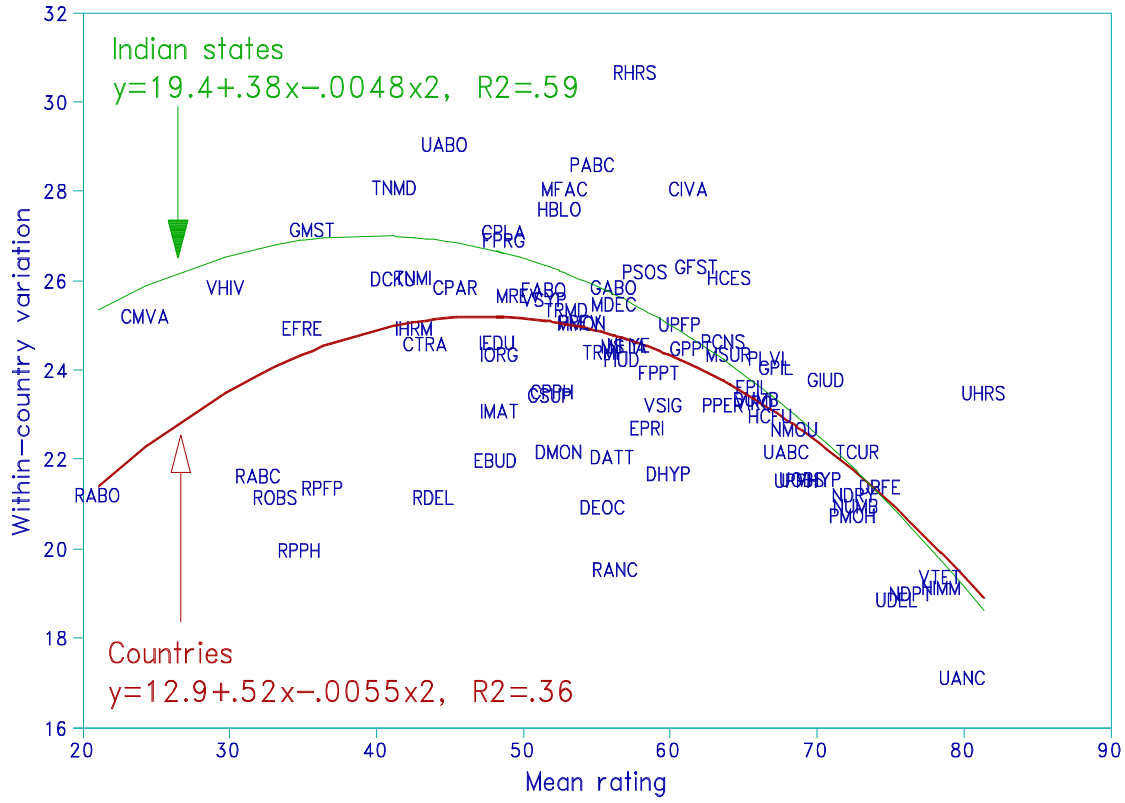
For ratings of Indian states, variation by rater position appears somewhat greater. (Note however the small number of raters in some groups, especially nurses and midwives.) Among these raters, it is outsiders without MDs more than outside MDs who give lower than average ratings. The ratings of outsiders without MDs are particularly low, relative to ratings of others, in the areas of access, care received, health promotion, and resources. In contrast, administrator MDs give higher ratings than others on policy and support services especially, as well as on family planning and various other areas.

Interpretations of these variations by rater position are certainly possible. It is important to recall, however, that overall the variations across country raters are usually small and often not significant. For raters of Indian states, significant effects are more likely but also less interpretable. Why there should be wider variation across Indian raters is not obvious and will require further discussion below.

Table B5. Adjusted deviations from grand mean for ratings of item groups, by rater position (adjusted with country mean as a covariate)

	Administrators		*Physi-cians*	*Nurses, midwives*	*Outsiders*		*(Grand*
Item group	*MDs*	*non-MDs*	*MDs*	*non-MDs*	*MDs*	*non-MDs*	*mean)*
Countries							
Health center capacity	0.6	1.2	-0.4	2.1	-3.9	1.7	(46.6)
Hospital capacity	1.8	1.3	-0.3	1.1	-2.2	-0.3	(61.0)
Rural access	2.0	-2.2	0.3	0.2	-0.9	-0.9	(38.4)
Urban access	5.2	-3.0	0.3	2.8	-2.8	-3.4	(67.3)
Antenatal care	0.7	4.1	-0.9	2.6	-2.5	-0.3	(59.6)
Delivery care	1.3	3.4	-1.6	3.6	-3.2	0.6	(57.8)
Newborn care	1.7	1.6	-0.8	3.2	-3.0	-0.2	(71.0)
Fam. plan. at centers	1.5	2.8	-1.3	3.2	-2.2	-0.3	(56.1)
Fam. plan. at hospitals	1.3	1.2	-0.9	1.4	-2.0	0.7	(58.2)
Policy	0.7	3.3	0.3	0.1	-2.4	-0.2	(62.1)
Resources	1.3	0.2	-0.1	1.3	-2.7	0.5	(46.8)
Monitoring	0.1	1.7	0.8	0.2	-3.8	1.0	(56.0)
Health promotion	-0.3	6.1	0.2	-2.0	-2.8	1.0	(46.3)
Training	1.9	3.8	-0.7	-2.2	-2.1	1.3	(52.9)
Minimum N	83	28	173	64	81	106	
Maximum N	120	56	243	90	122	161	
Indian states							
Health center capacity	3.6	-1.0	7.1	-13.7	-3.6	-1.8	(47.8)
Hospital capacity	2.2	-0.8	0.0	-5.9	-1.4	-0.4	(64.0)
Rural access	3.9	-3.5	1.0	2.4	1.3	-7.7	(52.8)
Urban access	1.3	-7.4	4.2	2.2	1.6	-4.2	(80.8)
Antenatal care	3.5	-3.2	4.9	-0.3	-0.7	-5.7	(62.1)
Delivery care	4.7	-3.2	0.5	1.5	-0.9	-4.3	(62.3)
Newborn care	4.7	-3.3	1.7	3.8	-2.9	-2.7	(75.4)
Fam. plan. at centers	6.8	-0.8	2.7	-2.1	-4.3	-3.4	(64.9)
Fam. plan. at hospitals	4.8	0.4	-0.3	-0.4	-2.2	-3.5	(80.8)
Policy	7.1	-2.2	-3.6	13.7	-4.0	-2.6	(67.2)
Resources	3.6	-2.6	0.6	4.8	0.1	-5.3	(67.7)
Monitoring	6.7	-2.2	3.1	4.6	-5.8	-2.3	(61.1)
Health promotion	5.6	0.7	3.7	-1.3	-2.9	-5.8	(58.6)
Training	4.9	1.2	-1.9	0.2	-2.3	-3.1	(60.5)
Minimum N	43	11	15	3	48	28	
Maximum N	65	15	28	5	68	46	

Figure B3. Rater variation within countries on specific items, plotted against mean ratings, and regression lines



Variation across items

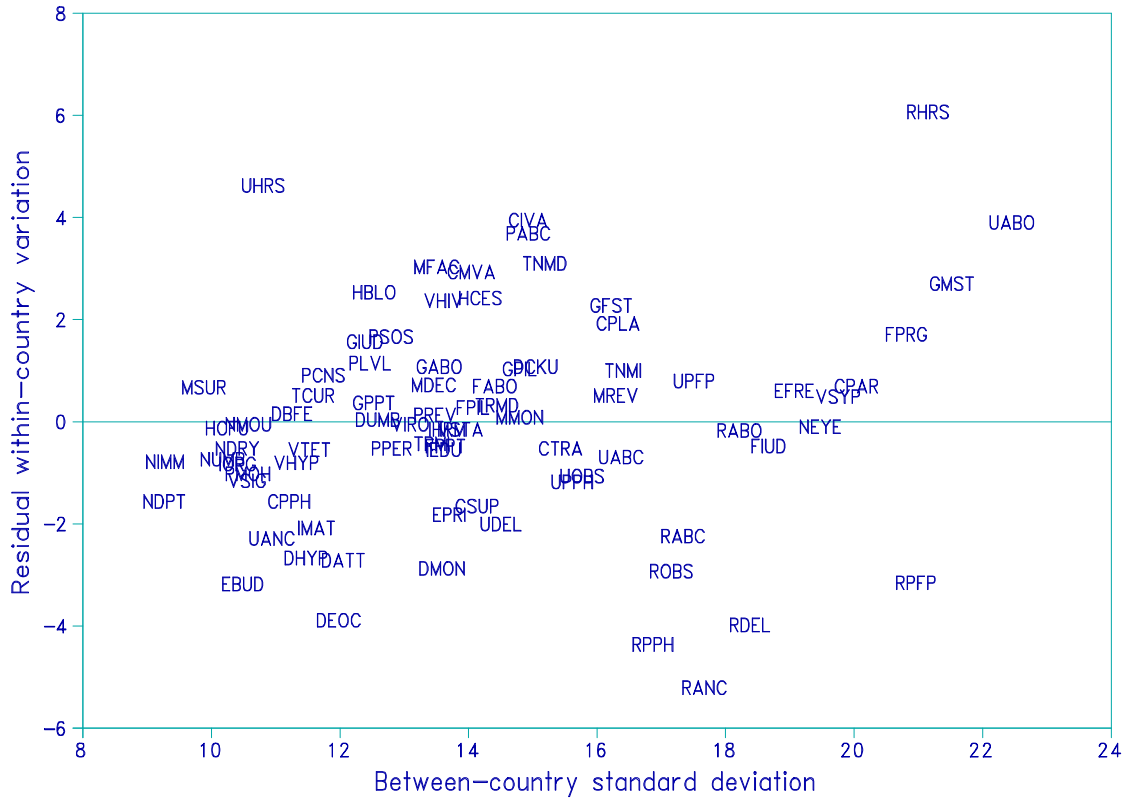
Given only small variations by category of rater, we consider all raters together to determine the extent to which they agree and disagree on different items.

Variation in ratings can be represented, across raters for one country, by the variance in their ratings. Averaging the variance across countries and taking the square root gives us an indicator of how much raters disagree overall on any specific item. Figure B3 plots this measure of variation against the mean rating for each item. Items are represented by short codes that can be found in Appendix A. (These estimates of variation leave out the individual Indian states, but India as a whole is included, represented by weighted averages across the states.)

Variation in ratings on one item is related to the mean rating for that item, being higher if the mean is approximately 50-60 than if the mean is above or below this range. This is a typical relationship for scales that resemble percentages (e.g., Tukey 1977), which have constrained minima and maxima. A quadratic regression line estimated from these data is shown in the figure; it has an R^2 of 0.36. A regression line for Indian states is also shown; it has essentially the same shape.

Taking residuals from the country regression line allows us to make fairer comparisons across items of rater agreement and disagreement. These residuals are plotted in Figure B4 against the between-

Figure B4. Residual within-country variation from regression line by item, plotted against between-country standard deviation



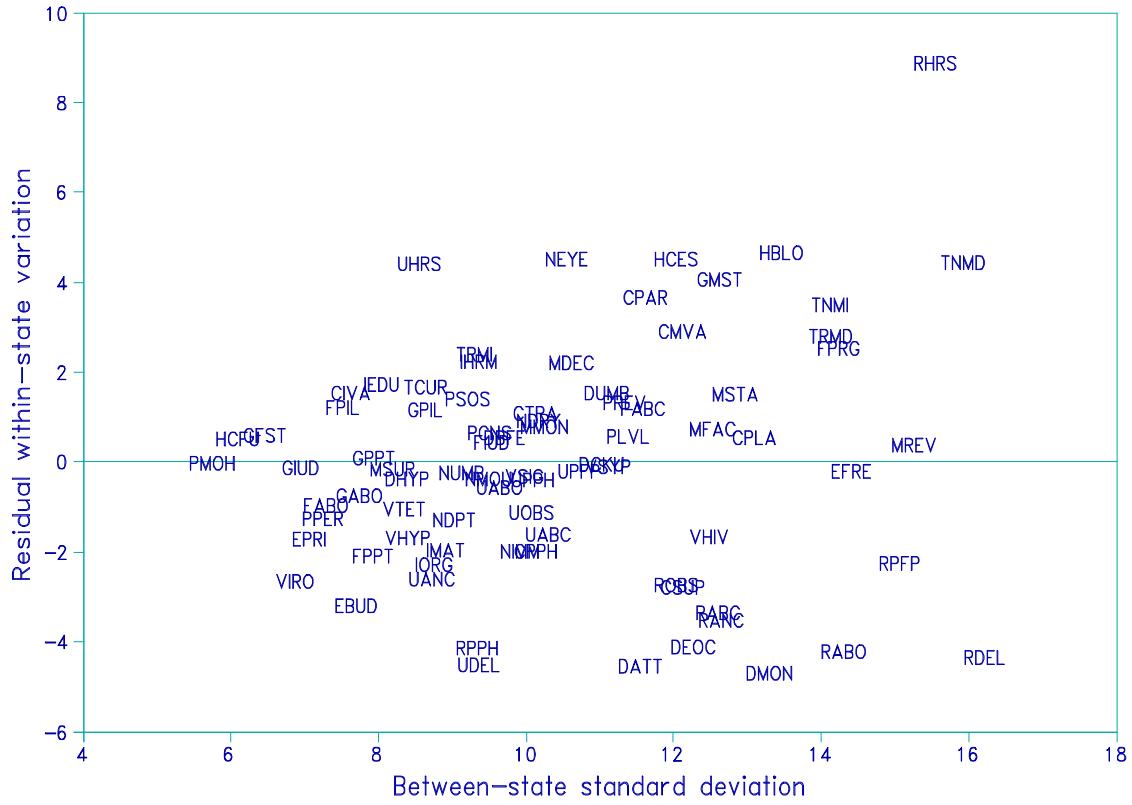
country standard deviations. A smaller residual is preferable, indicating that raters tend to agree among themselves, but a greater between-country standard deviation is also preferable, indicating that a given item has a greater ability to distinguish among countries.

By these criteria, the majority of the items on rural access (in the lower right quadrant of the figure) are particularly good items: raters tend agree in rating specific countries, and, at the same time, they recognize relatively greater distinctions among countries than on other items.

Four of the six items having to do with delivery care--emergency care, monitoring labor, a trained attendant, and checking for hypertension--also evoke relatively little disagreement among raters. Compared to the items on rural access, however, these items produce somewhat less variance across countries. One other item on which raters agree but on which countries do not differ as greatly is having adequate resources.

Although the access items as a group appear to be among the most satisfactory, the opposite is true in one case. Having access to 24-hour hospitals is problematic. Raters disagree more about these items (both rural and urban) than about any other items, and the urban item in particular does not provide much country contrast. These items, about whether "pregnant women have adequate access to district hospitals that are open 24 hours/day," may have been subject to alternative interpretations, some raters focusing on "adequate access" and others on whether the hospitals that exist are open all day and night.

Figure B5. Residual within-state variation from regression line for Indian states by item, plotted against between-state standard deviation



That said, the disadvantages of these items are relative; we have no absolute criteria for deciding how serious rater disagreement is.

Other items that evoke relatively more disagreement have to do with health center capacity to administer antibiotics intravenously, policy on treatment of abortion complications, and urban availability of abortions. Of these, only the last distinguishes relatively well among countries.

At least as notable as the specific items producing rater disagreement, however, is the fact that no group of items has consistently high residual variation. At least some items in each group elicit ratings that are relatively consistent across raters.

An alternative way to look at the levels of agreement and disagreement is to consider the widths of confidence intervals based on rating variance. Assume first that a typical country had exactly 16 raters (the median for all countries). The 95-percent confidence interval for the country score on the item with a median variance would be 23 points wide. The smallest confidence interval would be 15 points wide and the largest 28 points, but half of all confidence intervals would be between 20 and 24 points wide. Confidence intervals would be smaller for indices constructed by averaging or summing ratings.

Figure B5 uses the Indian state data in a graph similar to the preceding one. (Means and deviations are not weighted by population size in this graph.) Many of the details in this figure do not

match those in the previous figure, but in four main features the figures do agree: the majority of rural access items produce relatively good agreement among raters and relatively good discrimination between states; some items on care at delivery also produce good agreement but discriminate among the states less well; the resources item produces good agreement but poor discrimination; and items on access to 24-hour hospitals produce large disagreement.

Conclusion

We have investigated variations among raters of maternal and neonatal health services to see what clues this might provide to possible bias or inaccuracy in their ratings. We find that raters do differ and can disagree when rating the same country. But the differences are not large and the disagreements have some coherence, suggesting that the ratings are not arbitrary but probably have some basis.

Half of the raters skipped at least one item in the questionnaire. Those with the least national experience in maternal health were more likely to skip an item. On the other hand, MDs were less likely to skip any item. These contrasts are what one would expect if raters were judicious, answering questions on which they had some knowledge and skipping other questions.

Raters for Indian states follow these patterns somewhat less consistently. This may be simply a function of a smaller sample and greater sample variability, or there may be specific reasons having to do with the state programs (one is suggested below), or the Indian ratings may in fact have been somewhat less carefully produced. Unlike country raters, who were selected by local consultants supervised by the Futures Group International, raters for Indian states were selected in a three-step process. An Indian institution was first selected, which selected state consultants, who selected raters. The Futures Group later rejected ratings from one state (which are not included in the data) because they appeared to be fraudulent, a circumstance that did not occur for country ratings. The Indian data were scrutinized as carefully as the country data on receipt by the Futures Group but may nevertheless be of slightly lower quality.

When ratings are compared for the same countries (by controlling for mean country rating in an analysis of covariance), neither the number of items skipped nor rater experience nor rater position appears to lead to significant differences in the overall level of ratings. The lack of an effect of skipped items is important. It contradicts the argument that those who filled out the questionnaire more completely were simply guessing on many of the items. These conclusions appear to hold across regions; contrasts between raters categories did not differ by region. This finding could have one important implication. If it can be said that any one of the categories of raters, such as outside MDs, applies consistent standards across regions, this finding would imply that the other categories of raters are similar consistent across regions.

Despite the lack of results for overall rating level, ratings for item groups are sometimes significantly different by rater position. MDs who are outside the health program give lower ratings on all item groups, the difference being occasionally significant. The implication appears to be that outside observers who are fully aware of both the intricacies and the potential of health services but not deeply involved in providing them may be able to be more critical. These outside MDs, nevertheless, never rate any group of items more than 5 points below the ratings from other experts.

Policy and support services may be a relatively contentious area, where different categories of raters do not agree. No hard and fast standards exist in such areas as health promotion, so it is

understandable that people with different expertise, such as outside MDs, may provide different ratings from, say, administrators without MDs. In the cases of health promotion and training, however, it is important to note that nurses and midwives have slightly more negative evaluations than administrators, about as negative as those of outside MDs.

To interpret these results, one might hypothesize that raters give more positive ratings in their specific areas of expertise, as long as they are direct contributors to the program in that area. This need not be a deliberate attempt to mislead, but could simply reflect the individual's perspective and a natural psychological tendency to focus more on what is working than what is not.

This interpretation might account for various variations in ratings. Thus administrators with MDs (who are likely to have some connection with an urban hospital or clinic) rate urban access to services and hospital capacity more positively; administrators who are not MDs rate policy, health promotion, and training (areas that are more likely than medical interventions to reflect their expertise) more highly; nurses and midwives rate antenatal, delivery, and newborn care more positively, at the same time rating health promotion and training (not their expertise but activities they rely on) more negatively; and outside MDs, whose contribution to a program is less direct, rate everything more negatively. This is not a complete explanation, of course, because it fails to account for a few variations. It is important to note also that even these variations, which are the largest in the set, involve only a few points on 100-point scales, are often not significant, and should not be overinterpreted. Nevertheless this possible explanation indicates the appropriateness, for this exercise, of having used a variety of raters of different expertise and position.

For raters of Indian states, the contrasts by rater position are more likely to be significant and are somewhat greater than those for countries, though still mostly small. They are also somewhat less interpretable. Again this may simply be a function of the smaller number of raters or of looser controls over the rating process.

One contrast is in line with the previous speculations. Outsiders give generally lower ratings than administrators or providers. In each of these three groups, however, MDs give more positive ratings than non-MDs. This might be accounted for if we hypothesize that Indian state programs, in contrast to typical programs in other developing countries, are somewhat more medicalized, more directly under the control of medical doctors who exercise more intrusive authority down to lower levels of organization. Such a pattern of authority might, under the earlier hypothesis, lead MDs to provide more positive ratings than non-MDs.

We focused, finally, on differences in ratings across all raters regardless of rater position, to determine substantive areas on which raters are more and less likely to disagree. One striking result of this analysis was the broad similarity between results for country ratings and for ratings of Indian states. In both cases, ratings from different raters are more likely to be close together when the ratings are extreme ratings, whether high or low, than when they are intermediate. In both cases, raters agree most and disagree most on largely the same sets of items. They agree most on ratings of rural access, on several aspects of care at delivery, and on budget adequacy. They agree least on ratings of access to 24-hour hospitals. While the Indian ratings are possibly of lower quality, they still produce variations similar to those in country data.

The variability in rater reaction to different items is not unexpected and validates the approach of using multiple items in assessing different aspects of services. Similarly, the differences between raters,

while generally small, support the strategy of relying on multiple raters per country with different backgrounds. The contrasts described do not appear arbitrary but are generally small and usually interpretable, suggesting that raters were not making arbitrary judgments but were reflecting some common social reality--whether the actual state of services or the expert consensus about them--filtered through slightly different perspectives. With other data on raters, such as sex, age, and nationality, further contrasts might have been identified, but with the data available no substantial inconsistencies appear that would call into question the face validity of the ratings.

References

Bulatao, Rodolfo A. 1986. Dimensions of family planning effort. Paper presented at the Seminar on the Study of the Program Effort and Performance of Family Planning Services in Thailand, Bangkok.

Bulatao, Rodolfo A., and Laura B. Shrestha. 1996. Key indicators for reproductive health projects. World Bank, Washington, D.C.

Campbell, Oona, M. Koblinsky, and P. Taylor. 1995. Off to a rapid start: Appraising maternal mortality and services. *International Journal of Gynecology & Obstetrics* 48(Supplement):S33-S52.

Graham, Wendy J., and Véronique G. A. Filippi. 1994. Monitoring maternal health goals: How well do the indicators perform? *Maternal and Child Epidemiology Unit Publication No. 5.* London: London School of Hygiene and Tropical Medicine.

Koblinsky, Marge, Katie McLaurin, Pauline Russell-Brown, and Pamina Gorbach. 1995. *Indicators for Reproductive Health Program Evaluation: Final Report of the Subcommittee on Safe Pregnancy.* Chapel Hill, N.C.: EVALUATION Project, Carolina Population Center.

Maine, Deborah. 1991. *Safe Motherhood Programs: Options and Issues.* New York: Columbia University Center for Population and Family Health.

Maine, Deborah, Murat Z. Akalin, Victoria M. Ward, and Angela Kamara. 1997. *The Design and Evaluation of Maternal Health Programs.* New York: Columbia University Center for Population and Family Health.

Mauldin, W. Parker, and Bernard Berelson. 1978. Conditions of fertility decline in developing countries. *Studies in Family Planning* 9:89-147.

Mauldin, W. Parker, and Robert J. Lapham. 1985. Measuring family planning program effort in LDCs: 1972 and 1982. Pp. 1 40 in Nancy Birdsall, ed., *The Effects of Family Planning Programs on Fertility in the Developing World.* World Bank Staff Working Paper 677. Washington, D.C.: World Bank.

Mauldin, W. Parker, and John A. Ross. 1991. Family planning programs: Efforts and results, 1982-89. *Studies in Family Planning* 22(6):350-367.

Murgueytio, Patricio, Lucia Merino, and John Stover. 1997. The AIDS policy environment score: Measuring the degree to which the policy environment in Central America supports effective policies and programs for HIV/AIDS prevention, treatment, and care. Futures Group International, Glastonbury, Conn.

National Research Council. 1997. *Reproductive Health in Developing Countries: Expanding Dimensions, Building Solutions.* Amy O. Tsui, Judith N. Wasserheit, and John G. Haaga, eds. Washington, D.C.: National Academy Press.

Nunez, Cesar A., Patricio Murgueytio, and John Stover. 1998. The AIDS policy environment score: Measuring the degree to which the policy environment supports effective policies and programs for HIV/AIDS prevention, treatment and care; 1998 follow-up results. Futures Group International, Guatemala City.

Pham, Bich San, John A. Ross, Lan Phuong Nguyen, and Duc Vinh Nguyen. 1999. Measuring family

planning program effort at the provincial level: A Vietnam application. *International Family Planning Perspectives* 25(1):4-9.

Prevention of Maternal Mortality Network, The. 1995. Situation analyses of emergency obstetric care: Examples from eleven operations research projects in West Africa. *Social Science and Medicine* 40(5):657-667.

Ross, John A., and Katharine Cooper-Arnold. 2000. Effort scores for family planning programs: An alternative approach. Futures Group International, Glastonbury, Conn.

Ross, John A., and W. Parker Mauldin. 1996. Family planning programs: Efforts and results, 1972-94. *Studies in Family Planning* 27(3):137-147.

Schultz, T. Paul. 1994. Human capital, family planning, and their effects on population growth. *American Economic Review* 84(2):255-260.

Stover, John, Joel Rehnstrom, and Bernhard Schwartlander. 1999. Measuring the level of effort in the national and international response to HIV/AIDS: The AIDS program effort index (API). Paper presented at the XIIIth International AIDS Conference, Durban, South Africa, 9-13 July.

Tukey, John W. 1977. *Exploratory Data Analysis.* Reading, Mass.: Addison-Wesley Publishing Company.

United Nations Children's Fund (UNICEF), World Health Organization, and United Nations Population Fund. 1997. *Guidelines for Monitoring the Availability and Use of Obstetric Services.* New York.

World Health Organization. 1993. *Preventing Prolonged Labour: A Practical Guide.* Geneva.

World Health Organization. 1994. Indicators to monitor maternal health goals: Report of a technical working group, Geneva, 8-12 November 1993. Geneva.

World Health Organization. 1997. Mother-Baby Package Costing Spreadsheet. World Health Organization, Geneva.

STATISTICAL ADDENDUM

The MNPI questionnaire has 13 components that contain 81 items.
Within each component the items are averaged to produce a component score,
and these are converted to a 0-100 scale. A "total" score is also produced;
it is simply the mean for the 13 components, giving each one equal weight.

These tables give the results for all countries, for 1999 and 1996.

Statistical Addendum

1999 Ratings by Country, Region, and Component

(All ratings on a 0-100 scale)

	1	2	3	4	5	6	7	8	9	10	11	12	13	
	Capacities of health centers	Capacities of district hospitals	Percentage with access	Care at antenatal visits	Care at delivery	Care for newborns	Family planning at health centers	Family planning at district hospitals	Policies toward safe pregnancy	Resources	Information, education	Training arrangements	Monitoring, evaluation	**Total Score (Mean)**
East and Southeast Asia														
Cambodia	40.0	31.7	33.0	51.2	51.0	68.7	34.5	46.1	64.8	54.8	70.0	51.4	64.5	**50.9**
China	55.7	66.7	75.4	49.2	71.9	74.4	67.9	80.5	80.0	30.0	59.2	64.3	75.6	**65.4**
Indonesia	55.7	69.7	53.3	53.7	57.2	70.9	65.5	66.9	70.5	51.2	59.2	57.5	60.0	**60.9**
Myanmar	50.1	80.6	58.6	67.5	69.5	75.6	49.5	60.5	60.2	54.0	57.5	64.7	66.4	**62.7**
Philippines	24.4	40.7	68.9	51.5	59.5	73.0	59.3	61.5	56.1	53.6	60.0	61.0	59.3	**56.1**
Vietnam	32.1	65.0	73.9	63.2	65.8	73.5	59.3	73.6	67.6	51.1	65.2	64.2	67.2	**63.2**
Means for Region	**43.0**	**59.1**	**60.5**	**56.0**	**62.5**	**72.7**	**56.0**	**64.8**	**66.6**	**49.1**	**61.8**	**60.5**	**65.5**	**59.9**
South Asia														
Bangladesh	46.7	59.3	30.9	51.1	40.0	54.6	61.8	70.9	56.9	44.8	46.1	37.6	39.6	**49.3**
India	48.6	64.4	55.3	62.2	60.3	73.4	64.8	79.7	67.0	67.4	58.8	61.6	60.3	**63.4**
Nepal	23.7	45.6	17.3	42.1	38.3	59.0	36.8	54.8	60.7	38.6	44.6	33.4	42.2	**41.3**
Pakistan	30.1	40.6	24.6	31.7	28.6	45.6	29.2	40.8	33.6	37.8	22.9	22.3	28.1	**32.0**
Means for Region	**37.3**	**52.5**	**32.1**	**46.8**	**41.8**	**58.1**	**48.1**	**61.6**	**54.5**	**47.2**	**43.1**	**38.7**	**42.6**	**46.5**
Latin America and the Caribbean														
Bolivia	50.5	53.7	40.0	53.1	55.4	67.4	47.2	46.2	54.2	55.5	40.8	58.1	59.8	**52.4**
Brazil	53.0	74.0	64.2	64.7	64.3	77.7	54.7	63.7	62.3	65.3	54.0	54.8	63.7	**62.8**
Dominican Republi	55.5	64.3	70.1	67.1	69.4	83.3	58.0	54.8	65.0	51.4	47.5	64.6	63.1	**62.6**
Ecuador	43.5	64.0	53.4	69.2	59.2	76.7	57.9	58.7	57.0	50.2	42.9	49.3	56.9	**56.8**
El Salvador	47.3	63.3	48.0	63.9	50.5	72.8	60.0	65.6	48.1	36.7	30.8	64.0	57.7	**54.5**
Guatemala	48.3	61.3	42.5	60.9	54.2	69.4	45.1	43.2	47.6	49.7	25.9	53.7	51.9	**50.3**
Haiti	37.4	55.9	31.7	52.0	45.2	53.2	38.2	48.0	44.5	41.4	30.3	37.1	33.2	**42.2**
Honduras	45.0	64.6	49.5	60.8	55.2	72.3	48.2	56.0	63.8	50.9	44.0	61.7	56.1	**56.0**
Jamaica	45.6	72.1	82.9	75.3	75.8	88.0	79.9	72.2	71.0	44.5	43.2	68.2	76.2	**68.9**
Mexico	48.9	68.3	67.3	54.2	63.1	78.6	64.3	78.1	47.5	44.8	46.3	61.0	53.5	**59.7**
Nicaragua	46.1	62.4	50.0	65.6	56.5	74.5	59.0	60.8	48.1	43.6	35.5	56.6	52.9	**54.7**
Paraguay	49.3	60.0	56.7	64.3	63.0	70.2	61.2	64.1	62.7	42.5	44.1	53.9	43.7	**56.6**
Peru	65.9	69.2	71.2	74.0	64.3	73.5	76.9	79.4	65.7	54.6	44.7	68.3	63.8	**67.0**
Means for Region	**49.0**	**64.1**	**56.0**	**63.5**	**59.7**	**73.7**	**57.7**	**60.8**	**56.7**	**48.6**	**40.8**	**57.8**	**56.3**	**57.3**
Middle East and North Africa														
Algeria	55.4	79.3	66.5	50.7	64.9	73.3	50.8	59.3	43.5	50.4	40.6	51.8	50.3	**56.7**
Egypt	47.5	68.7	74.3	60.3	63.1	79.3	70.8	62.2	71.1	64.0	47.0	54.8	59.8	**63.3**
Iran	60.8	82.2	80.8	72.8	78.7	86.8	87.1	83.3	75.0	67.0	65.8	61.4	66.0	**74.4**
W. Bank	43.1	58.7	73.4	50.0	62.5	79.3	45.9	38.9	54.8	62.5	47.5	57.4	50.4	**55.7**
Yemen	17.2	28.0	29.3	36.9	29.2	41.3	41.6	44.0	48.7	35.2	35.3	26.3	30.1	**34.1**
Means for Region	**44.8**	**63.4**	**64.9**	**54.1**	**59.7**	**72.0**	**59.2**	**57.5**	**58.6**	**55.8**	**47.2**	**50.3**	**51.3**	**56.8**

Non-Francophone and Sub-Saharan Africa														
Angola	53.7	58.2	36.0	67.8	67.7	69.1	62.6	55.2	62.4	39.1	54.0	64.4	66.4	**58.2**
Ethiopia	47.1	60.4	27.3	51.3	41.1	59.2	59.8	53.6	57.1	38.8	41.2	39.4	48.4	**48.1**
Ghana	58.9	77.5	54.1	65.1	62.3	78.5	73.9	72.7	85.5	53.7	48.8	54.7	68.3	**65.7**
Kenya	35.4	54.6	40.7	46.7	45.0	64.4	64.6	71.4	55.0	33.8	27.5	33.2	38.9	**47.0**
Malawi	55.5	74.0	54.0	63.7	69.6	72.6	65.4	68.1	76.1	62.1	59.6	55.3	53.0	**63.8**
Mozambique	46.2	46.7	42.1	48.4	52.9	66.1	50.2	47.3	69.3	34.3	42.8	55.2	68.3	**51.5**
Nigeria	42.4	54.0	40.9	57.3	53.4	63.7	46.6	47.8	56.6	37.6	51.5	39.3	43.8	**48.8**
South Africa	59.2	62.9	73.2	66.8	58.5	76.3	62.0	56.9	64.7	64.7	42.5	57.4	54.5	**61.5**
Sudan	31.0	49.7	52.1	47.3	55.2	67.1	52.7	38.6	57.3	41.3	43.8	68.8	58.3	**51.0**
Tanzania	41.5	65.8	47.3	54.7	48.6	62.0	52.8	57.2	65.3	46.2	57.7	61.3	-	**50.8**
Uganda	49.6	64.4	41.3	58.2	54.8	67.3	58.4	62.2	68.7	54.8	62.3	56.1	59.0	**58.3**
Zambia	32.9	59.0	37.4	48.4	47.7	61.0	56.7	59.2	58.4	32.8	35.2	29.3	46.4	**46.5**
Zimbabwe	49.0	67.7	65.9	69.0	69.0	85.0	63.4	63.3	71.0	43.1	55.0	61.5	72.8	**64.3**
Means for Region	**46.3**	**61.1**	**47.1**	**57.3**	**55.8**	**68.7**	**59.2**	**58.0**	**65.2**	**44.8**	**47.8**	**52.0**	**52.2**	**55.0**
Francophone and Sub-Saharan Africa														
Benin	58.5	63.6	49.2	63.1	62.7	77.5	59.2	60.6	66.1	40.6	39.6	38.6	58.2	**56.7**
Congo	44.4	55.9	52.1	60.5	57.4	74.9	39.6	35.1	65.0	37.9	44.4	45.6	36.6	**50.0**
Congo, Dem. Rep.	51.2	63.6	38.2	66.7	60.3	71.3	36.6	39.3	45.5	35.2	32.1	48.4	41.0	**48.4**
Guinea	45.4	59.2	37.2	68.6	64.1	78.8	59.5	48.0	80.2	44.7	58.4	54.3	62.1	**58.5**
Madagascar	54.8	51.5	48.1	71.9	68.7	79.1	64.7	61.7	68.5	42.6	46.3	45.1	68.0	**59.3**
Mali	64.9	62.5	42.2	70.7	69.1	74.6	66.1	63.6	73.8	39.6	60.0	68.3	72.9	**63.7**
Rwanda	47.5	66.3	43.9	59.4	45.4	68.4	34.1	53.1	67.9	40.1	45.9	48.2	50.2	**51.6**
Senegal	63.3	63.2	39.6	67.5	67.3	75.1	72.2	62.9	72.8	38.9	38.6	49.3	50.3	**58.5**
Means for Region	**53.7**	**60.7**	**43.8**	**66.0**	**61.9**	**75.0**	**54.0**	**53.0**	**67.5**	**40.0**	**45.7**	**49.7**	**54.9**	**55.8**
Grand Mean of All Countries	**46.9**	**61.1**	**51.1**	**59.0**	**57.9**	**71.0**	**56.7**	**59.0**	**61.9**	**46.8**	**46.9**	**53.0**	**54.5**	**55.8**
Indian States														
Andhra Pradesh	42.1	78.8	65.5	72.7	66.2	81.5	67.0	91.0	77.3	78.8	79.1	81.3	73.1	**73.4**
Assam	42.6	67.8	47.2	47.2	48.6	59.7	52.5	66.9	58.8	61.7	52.5	63.7	46.4	**55.0**
Gujrat	64.2	61.6	65.3	65.9	70.0	82.9	74.7	87.4	73.2	78.5	64.8	75.7	68.6	**71.7**
Haryana	56.7	80.0	65.0	60.0	65.3	79.7	64.8	87.7	68.6	64.1	58.1	68.7	65.6	**68.0**
Karnataka	37.0	47.5	60.0	56.8	50.5	72.6	59.7	76.6	56.7	62.2	50.3	44.2	52.2	**55.9**
Kerala	41.0	61.0	74.5	70.3	78.4	86.0	69.3	81.0	66.7	72.2	58.5	47.0	61.3	**66.7**
Madhya Pradesh	48.4	69.2	57.7	70.0	59.4	74.8	67.6	86.0	75.7	61.8	62.5	61.1	53.1	**65.2**
Maharastra	61.2	75.0	63.4	67.5	67.9	81.9	67.3	77.8	71.8	72.1	64.4	73.0	76.7	**70.7**
Orissa	52.7	48.2	40.5	58.8	63.5	80.0	71.5	82.1	59.7	61.3	60.5	53.8	53.7	**60.5**
Punjab	48.6	67.7	60.7	55.8	56.5	74.9	67.8	88.2	65.7	59.5	50.8	67.1	66.4	**63.8**
Rajasthan	44.1	58.6	43.5	47.7	52.3	64.6	57.7	73.0	66.5	57.9	54.7	53.5	58.0	**56.3**
Tamil Nadu	47.0	67.9	78.6	70.2	77.1	82.9	66.7	80.0	78.3	80.4	55.6	52.1	72.8	**70.0**
Uttar Pradesh	45.2	59.4	41.9	56.7	50.4	61.4	58.6	72.6	59.9	63.5	52.6	59.2	54.8	**56.6**
West Bengal	48.1	59.3	51.3	60.3	54.6	67.3	63.6	80.0	58.0	64.0	53.0	54.4	44.7	**58.4**
India *	**48.6**	**64.4**	**55.3**	**62.2**	**60.3**	**73.4**	**64.8**	**79.7**	**67.0**	**67.4**	**58.8**	**61.6**	**60.3**	**63.4**

*States weighted by population

Statistical Addendum

1996 Ratings by Country, Region, and Component

(All ratings on a 0-100 scale)

	1	2	3	4	5	6	7	8	9	10	11	12	13	
	Capacities of health centers	Capacities of district hospitals	Percentage with access	Care at antenatal visits	Care at delivery	Care for newborns	Family planning at health centers	Family planning at district hospitals	Policies toward safe pregnancy	Resources	Information, education	Training arrangements	Monitoring, evaluation	**Total Score (Mean)**
East and Southeast Asia														
Cambodia	14.2	13.7	16.0	20.2	17.6	34.2	11.6	18.3	29.5	27.6	30.9	19.0	25.8	**21.4**
China	41.2	54.4	68.6	34.4	60.5	65.9	60.2	79.1	73.8	22.2	44.2	57.0	68.6	**56.2**
Indonesia	40.0	52.5	45.2	39.5	42.8	63.3	64.5	66.7	55.5	45.2	46.3	45.0	50.0	**50.5**
Myanmar	43.4	75.6	50.1	57.8	58.3	68.0	34.3	52.0	52.6	51.4	50.0	56.3	56.4	**54.3**
Philippines	15.1	34.0	63.0	41.9	46.8	63.6	52.0	54.6	46.1	51.2	48.6	58.8	53.6	**48.4**
Vietnam	20.5	45.0	61.5	39.6	45.4	54.9	41.0	56.4	52.9	37.8	46.5	46.8	54.2	**46.3**
Means for Region	**29.0**	**45.9**	**50.7**	**38.9**	**45.2**	**58.3**	**43.9**	**54.5**	**51.7**	**39.2**	**44.4**	**47.2**	**51.4**	**46.2**
South Asia														
Bangladesh	33.6	39.3	24.7	35.9	26.8	45.9	52.7	63.0	41.0	35.2	26.1	24.1	24.6	**36.4**
India	32.5	46.4	44.6	43.3	40.5	57.2	49.4	65.4	51.9	52.6	41.3	45.4	44.0	**47.3**
Nepal	8.6	20.5	12.0	26.3	21.4	42.2	22.8	36.1	32.5	22.3	20.0	16.9	20.8	**23.3**
Pakistan	23.2	32.2	22.2	24.2	21.2	36.9	20.1	32.2	24.9	31.1	15.0	18.0	20.3	**24.7**
Means for Region	**24.5**	**34.6**	**25.9**	**32.4**	**27.5**	**45.6**	**36.2**	**49.2**	**37.6**	**35.3**	**25.6**	**26.1**	**27.4**	**32.9**
Latin America and the Caribbean														
Bolivia	37.7	42.0	36.9	40.6	46.2	59.8	40.4	40.0	51.6	51.3	38.7	53.6	60.6	**46.1**
Brazil	41.4	60.0	58.3	53.0	55.1	71.0	40.9	52.6	49.7	57.3	45.5	46.4	55.3	**52.8**
Dominican Republic	48.7	61.0	69.9	61.3	64.6	84.1	55.4	52.8	59.8	47.0	47.7	60.0	56.9	**59.2**
Ecuador	33.0	55.7	47.9	56.0	48.9	71.6	51.2	50.8	44.3	43.1	31.0	38.7	43.9	**47.4**
El Salvador	41.1	37.5	47.7	41.7	31.4	53.8	37.1	41.0	27.8	17.1	21.4	46.9	43.8	**37.6**
Guatemala	23.0	47.9	29.3	34.4	32.9	51.5	27.7	29.1	29.3	36.9	14.8	33.8	30.9	**32.4**
Haiti	28.6	43.9	26.0	42.9	34.3	43.7	26.1	37.8	37.1	38.1	22.2	32.4	27.5	**33.9**
Honduras	33.8	53.5	42.6	50.0	44.6	63.3	40.8	49.3	52.4	45.9	35.6	56.6	46.9	**47.3**
Jamaica	42.5	71.1	81.5	69.9	70.0	86.6	75.4	71.4	69.2	49.6	42.1	65.9	73.0	**66.8**
Mexico	38.7	56.4	61.9	37.3	45.2	66.0	60.7	68.5	38.8	36.5	32.7	49.8	43.1	**48.9**
Nicaragua	40.9	56.7	50.7	60.8	51.6	75.8	56.2	59.3	47.7	44.1	34.1	57.7	50.5	**52.8**
Paraguay	37.5	46.7	49.6	54.2	52.3	64.4	53.5	54.4	54.1	37.5	32.6	47.5	36.6	**47.8**
Peru	33.5	33.3	51.4	44.6	39.5	51.5	46.7	51.7	40.4	34.6	20.6	49.3	40.1	**41.3**
Means for Region	**37.0**	**51.2**	**50.3**	**49.7**	**47.4**	**64.8**	**47.1**	**50.7**	**46.3**	**41.5**	**32.2**	**49.1**	**46.9**	**47.3**
Middle East and North Africa														
Algeria	55.4	78.7	67.1	49.7	62.3	73.7	50.8	59.3	43.5	48.1	40.0	54.6	57.3	**57.0**
Egypt	31.1	44.2	61.3	35.8	37.5	57.8	46.8	45.6	48.8	44.8	25.0	36.0	42.0	**42.8**
Iran	55.7	76.5	77.3	68.1	72.4	84.4	83.2	76.9	70.0	64.1	58.5	54.5	61.0	**69.4**
W. Bank	34.4	51.7	68.3	42.9	53.5	73.8	26.4	26.9	35.4	51.5	30.8	40.4	38.2	**44.2**
Yemen	12.8	23.6	27.0	33.1	23.0	35.6	32.6	36.7	34.4	26.0	20.0	18.0	21.9	**26.5**
Means for Region	**37.9**	**54.9**	**60.2**	**45.9**	**49.8**	**65.0**	**48.0**	**49.1**	**46.4**	**46.9**	**34.9**	**40.7**	**44.1**	**48.0**

Non-Francophone and Sub-Saharan Africa														
Angola	44.8	52.0	34.2	62.6	61.3	63.2	59.3	49.7	54.3	35.0	50.3	59.9	63.2	**53.1**
Ethiopia	44.1	54.9	22.1	48.6	35.9	54.5	49.2	45.3	51.0	36.6	34.4	35.1	42.4	**42.6**
Ghana	35.1	53.8	41.2	48.8	42.0	65.0	55.0	57.0	67.1	37.5	33.8	37.3	52.3	**48.1**
Kenya	34.8	58.2	37.4	44.2	40.9	62.1	61.7	65.8	47.6	32.6	28.1	30.6	34.7	**44.5**
Malawi	44.7	64.0	46.9	57.9	58.8	72.1	58.9	55.7	66.1	58.0	37.5	44.0	39.4	**54.2**
Mozambique	31.9	29.1	30.8	33.5	38.5	53.2	33.1	32.3	52.9	23.4	28.3	42.0	50.5	**36.9**
Nigeria	29.7	39.5	33.5	43.1	37.1	49.4	32.2	35.9	42.2	29.4	36.8	28.2	35.7	**36.4**
South Africa	48.6	61.4	64.0	54.8	52.4	73.0	53.5	52.9	54.3	58.4	37.5	52.3	40.1	**54.1**
Sudan	19.3	34.6	45.3	32.4	40.3	54.1	35.5	28.3	44.6	34.2	29.4	60.3	46.0	**38.8**
Tanzania	31.6	56.9	41.8	49.1	39.0	55.1	43.5	45.3	55.5	38.5	49.7	53.3	-	**43.0**
Uganda	23.7	50.7	32.3	41.0	36.6	56.0	36.6	46.9	50.4	41.7	43.0	39.0	40.5	**41.4**
Zambia	27.1	51.0	33.2	41.9	38.7	53.5	45.8	48.6	49.3	33.2	28.0	24.7	36.0	**39.3**
Zimbabwe	44.8	63.6	61.1	60.8	59.8	79.5	57.2	53.1	62.9	44.6	48.1	55.7	66.2	**58.3**
Means for Region	**35.4**	**51.5**	**40.3**	**47.6**	**44.7**	**60.8**	**47.8**	**47.5**	**53.7**	**38.7**	**37.3**	**43.3**	**42.1**	**45.4**
Francophone and Sub-Saharan Africa														
Benin	47.8	50.8	43.7	56.3	53.1	69.4	45.9	51.8	52.2	33.4	33.3	34.3	50.7	**47.9**
Congo	44.6	49.7	51.8	56.2	51.5	72.4	40.7	35.0	62.2	39.5	48.2	48.2	39.7	**49.2**
Congo, Dem. Rep. Of	50.1	56.9	39.6	57.4	57.2	69.2	32.9	37.1	36.1	30.0	27.9	49.0	32.9	**44.3**
Guinea	27.1	39.2	24.8	53.0	46.5	71.1	42.8	32.0	64.3	35.7	44.4	38.7	46.5	**43.6**
Madagascar	34.1	35.2	36.3	50.2	50.2	66.7	51.6	48.5	38.0	37.2	17.0	17.3	35.7	**39.8**
Mali	37.4	40.4	29.5	49.2	47.7	60.3	51.6	49.5	61.0	29.2	47.0	54.8	55.8	**47.2**
Rwanda	27.1	41.2	30.9	37.8	27.9	55.6	23.0	38.4	55.8	30.6	27.9	34.8	31.9	**35.6**
Senegal	44.1	44.8	31.4	51.5	48.7	66.3	55.8	50.3	54.4	30.4	26.2	38.4	47.7	**45.4**
Means for Region	**39.0**	**44.8**	**36.0**	**51.4**	**47.8**	**66.4**	**43.0**	**42.8**	**53.0**	**33.2**	**34.0**	**39.4**	**42.6**	**44.1**
Grand Mean of All Countries	**35.0**	**48.6**	**44.4**	**46.3**	**45.1**	**61.7**	**45.4**	**48.7**	**49.3**	**39.2**	**35.1**	**43.0**	**43.6**	**45.0**
Indian States														
Andhra Pradesh	19.5	46.3	51.3	44.4	37.0	59.2	48.3	74.6	56.3	61.7	55.9	61.0	53.1	**51.4**
Assam	34.5	58.9	41.5	35.0	37.6	50.3	43.2	56.9	47.4	52.2	41.7	54.5	38.7	**45.6**
Gujrat	49.2	46.5	55.4	48.0	50.6	69.8	63.8	79.1	61.5	69.4	52.0	66.2	50.8	**58.7**
Haryana	34.5	55.4	54.0	41.7	44.2	63.8	51.2	76.4	54.3	50.3	40.4	49.7	46.9	**51.0**
Karnataka	21.0	31.0	50.9	38.2	33.5	59.1	45.2	61.3	42.5	47.0	35.8	30.2	37.4	**41.0**
Kerala	30.8	48.3	64.6	56.7	63.1	72.5	57.7	69.3	53.9	61.4	45.6	36.7	48.6	**54.6**
Madhya Pradesh	31.6	43.0	46.4	48.2	36.4	57.5	51.9	72.8	60.4	41.7	40.3	43.9	33.5	**46.7**
Maharastra	47.8	66.3	54.7	56.3	56.4	72.3	54.8	66.0	67.3	65.0	55.0	61.0	69.8	**61.0**
Orissa	28.5	28.1	27.3	34.5	36.8	54.5	47.2	61.8	34.9	43.3	32.5	30.3	28.3	**37.5**
Punjab	38.9	59.0	53.9	50.3	47.0	69.2	62.0	86.4	60.9	57.9	45.8	62.8	62.4	**58.2**
Rajasthan	31.8	40.8	32.4	31.9	33.5	47.7	42.2	57.0	52.8	44.1	40.5	38.1	42.4	**41.2**
Tamil Nadu	28.2	45.8	68.6	51.0	58.2	69.3	55.8	68.4	62.1	68.9	42.5	40.9	50.2	**54.6**
Uttar Pradesh	28.3	40.7	28.4	35.2	26.6	40.7	40.5	53.6	39.9	41.6	30.8	37.1	35.7	**36.8**
West Bengal	34.7	49.3	41.2	41.2	38.3	52.0	45.2	65.3	43.4	50.0	32.0	37.6	29.7	**43.1**
India *	**32.5**	**46.4**	**44.6**	**43.3**	**40.5**	**57.2**	**49.4**	**65.4**	**51.9**	**52.6**	**41.3**	**45.4**	**44.0**	**47.3**

*States weighted by population