Thirty-five years later: Long-term effects of the Matlab maternal and child health/family planning program on older women's well-being

Tania Barham[a,b], Brachel Champion[a,b], Andrew D. Foster[c,d], Jena D. Hamadani[e], Warren C. Jochem[f], Gisella Kagy[g], Randall Kuhn[h,1], Jane Menken[b,i,1], Abdur Razzaque[j], Elisabeth Dowling Root[k], and Patrick S. Turner[l]

[a]Department of Economics, University of Colorado Boulder, Boulder, CO 80302; [b]Institute of Behavioral Science, University of Colorado Boulder, Boulder, CO 80302; [c]Department of Economics, Brown University, Providence, RI 02903; [d]Department of Health Services, Policy and Practice, Brown University, Providence, RI 02903; [e]Maternal and Child Health Division, icddr,b, Dhaka 1212, Bangladesh; [f]School of Geography and Environmental Science, University of Southampton, Southampton SO17 1BJ, United Kingdom; [g]Department of Economics, Vassar College, Poughkeepsie, NY 12604; [h]Department of Community Health Sciences, UCLA Fielding School of Public Health, Los Angeles, CA 90095; [i]Department of Sociology, University of Colorado Boulder, Boulder, CO 80302; [j]Health System and Population Studies Division, icddr,b, Dhaka 1212, Bangladesh; [k]Department of Geography, The Ohio State University, Columbus, OH 43210; and [l]Wilson Sheehan Lab for Economic Opportunities, Department of Economics, University of Notre Dame, Notre Dame, IN 46556

Contributed by Jane Menken, May 9, 2021 (sent for review January 19, 2021; reviewed by Elizabeth Frankenberg and David A. Lam)

Family planning programs are believed to have substantial long-term benefits for women's health and well-being, yet few studies have established either extent or direction of long-term effects. The Matlab, Bangladesh, maternal and child health/family planning (MCH/FP) program afforded a 12-y period of well-documented differential access to services. We evaluate its impacts on women's lifetime fertility, adult health, and economic outcomes 35 y after program initiation. We followed 1,820 women who were of reproductive age during the differential access period (born 1938–1973) from 1978 to 2012 using prospectively collected data from the Matlab Health and Demographic Surveillance System and the 1996 and 2012 Matlab Health and Socioeconomic Surveys. We estimated intent-to-treat single-difference models comparing treatment and comparison area women. MCH/FP significantly increased contraceptive use, reduced completed fertility, lengthened birth intervals, and reduced age at last birth, but had no significant positive impacts on health or economic outcomes. Treatment area women had modestly poorer overall health (+0.07 SD) and respiratory health (+0.12 SD), and those born 1950–1961 had significantly higher body mass index (BMI) in 1996 (0.76 kg/m^2) and 2012 (0.57 kg/m^2); fewer were underweight in 1996, but more were overweight or obese in 2012. Overall, there was a +2.5 kg/m^2 secular increase in BMI. We found substantial changes in lifetime contraceptive and fertility behavior but no long-term health or economic benefits of the program. We observed modest negative health impacts that likely result from an accelerated nutritional transition among treated women, a transition that would, in an earlier context, have been beneficial.

family planning | long-term follow-up | health

The case for global scale-up of family planning programs rests, in part, on the potential long-term benefits of family planning programs for women's health and economic empowerment (1, 2). A counterpoint to this assumption suggests that smaller families may actually have negative consequences for women in societies where old-age support and women's status are tied to childbearing (3). Yet, few studies have established either the extent or direction of long-term effects.

Much of our existing understanding of family planning program effects comes from the Matlab maternal and child health/family planning (MCH/FP) program (4–6). The program was implemented by icddr,b (formerly known as the International Centre for Diarrhoeal Disease Research, Bangladesh) starting in 1977 in the rural Matlab area of Bangladesh. MCH/FP revolutionized the field by using a home-based delivery model, integrating family planning with mother-and-child health services, and collecting extensive data that facilitate evaluation. MCH/FP yielded immediate and enduring effects. While fertility levels were similar in treatment and comparison areas at baseline, the MCH/FP treatment area subsequently experienced increased contraceptive use, reduced fertility (6–8), and reductions in maternal and child mortality (9). By 1989, MCH/FP services were scaled up to the rest of Bangladesh, including the comparison area, creating a well-documented period of differential treatment exposure from 1977 to 1989. Using the 1996 Matlab Health and Socioeconomic Survey (MHSS1), Barham (5) assessed the medium-term effects of MCH/FP on beneficiary children in a representative sample of Matlab. Joshi and Schultz (4) found that adult women living in the treatment area compared to comparison area villages had a body mass index (BMI) more than 1 kg/m^2 higher and concomitant reductions in the proportion with BMI of <18 kg/m^2. They found no differences in self-rated health or self-reported activities of daily living (ADLs); however, the targeted

SOCIAL SCIENCES

ECONOMIC SCIENCES

Significance

Few studies have addressed links between family planning programs and long-term benefits for women's health and economic outcomes, especially in societies where old-age support and women's status are tied to childbearing and where smaller families may carry negative consequences for women. We analyzed the maternal and child health/family planning (MCH/FP) program, a highly effective intervention introduced in the rural Matlab subdistrict of Bangladesh in 1977 with a subsequent 12-y differential in service access. We found significant differences in lifetime contraceptive behavior and completed fertility among women born 1938–1973. We found few effects on later health or economic outcomes except for an association of MCH/FP with poorer overall health and poorer metabolic health among women born 1950–1961.

Author contributions: T.B., A.D.F., W.C.J., R.K., J.M., A.R., and E.D.R. designed research; T.B., J.D.H., W.C.J., G.K., R.K., J.M., A.R., E.D.R., and P.S.T. performed research; T.B., B.C., J.D.H., G.K., R.K., J.M., E.D.R., and P.S.T. analyzed data; and T.B., B.C., A.D.F., G.K., R.K., J.M., A.R., E.D.R., and P.S.T. wrote the paper.

Reviewers: E.F., The University of North Carolina at Chapel Hill; and D.A.L., University of Michigan–Ann Arbor.

The authors declare no competing interest.

See online for related content such as Commentaries.

[1]To whom correspondence may be addressed. Email: kuhn@ucla.edu or menken@colorado.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2101160118/-/DCSupplemental.

Published July 6, 2021.

women were still relatively young in 1996, and these measures of health typically show variation later in life.

Relatively few other studies have demonstrated long-term effects of family planning programs, in part due to the difficulties of longitudinal follow-up, biased self-selection into treatment, and the lack of appropriate comparison groups. Family planning may affect long-term well-being through a multiplicity of mediating pathways, including via effects on age at first birth, spacing between births, age at last birth, and completed fertility (2). Canning and Schultz (1) recently highlighted similar pathways of later benefit of family planning as potentially freeing up resources and women's time. Miller (10) showed that early access to family planning in Colombia led to delayed fertility, higher schooling, and greater labor force participation. Yet, to our knowledge, only Matlab offers the potential to apply a prospective intervention design to look at long-term health effects.

This study examined the consequences of the Matlab MCH/FP on women's lifetime fertility, adult health, and economic outcomes using the second Matlab Health and Socioeconomic Survey (MHSS2), conducted in 2012, ∼35 y after initial rollout of services. We focused on women who were of reproductive age during the period of differential treatment exposure—those born 1938–1973 (aged 40 y to 75 y in 2012). We measured impacts on three domains of health—metabolic, functional, and respiratory health—using a mix of directly observed physical tests (i.e., anthropometry, blood pressure, and grip strength) and indirectly diagnosed morbidity. The rich integrated database of the icddr,b Matlab Health and Demographic Surveillance System (HDSS) combined with MHSS1 and MHSS2 allowed us to address causal considerations related to assignment of intent-to-treat (ITT) status before program introduction, to selective attrition, and to potential confounders. We also examined take-up of modern contraception as well as underlying mechanisms related to later health, including family size, birth spacing, and BMI. In addition, we measured the effect of the program on economic outcomes such as consumption, savings, and employment.

Results

Study Design. The MCH/FP program was initiated in Matlab, a rural area of Bangladesh, in 1977 by icddr,b. Treatment and comparison areas were built into program design (11) and covered about 200,000 people in 149 villages, with the population split fairly evenly between the two areas (*SI Appendix*, section 1). The program was placed into geographically contiguous treatment and comparison areas (*SI Appendix*, Fig. S1). This block design was intended to reduce potential contamination of spillover into comparison groups that would occur in a randomized trial (12). Comparison villages were viewed as socially and economically similar (13). Blocks were chosen to balance average distance to transport and health infrastructure. We thus refer to the placement of this intervention as *quasi-random* and draw further support for our identification strategy from preprogram similarities between treatment and comparison areas (see *SI Appendix*, section 2). Family planning services began in October 1977, with mother-and-child health services phased in beginning in 1982. By 1989, MCH/FP services were scaled up to the rest of Bangladesh, including the comparison area, creating a well-documented period of differential treatment exposure from 1977 to 1989. Our analysis focused on all women born 1938–1973 and on three birth cohorts (1938–1949, 1950–1961, and 1962–1973) chosen by their age at first exposure to MCH/FP. In 1977, these cohorts were aged 28 y to 39 y, 16 y to 27 y, and 4 y to 15 y, respectively, so are considered as representing later, peak, and prereproductive life at the start of the program.

Data and Sample. Analysis used the extraordinarily rich data available for the Matlab study area and included three main data sources. First, we used the Matlab HDSS to identify our study population. This unique longitudinal demographic registration system recorded, within a short time of their occurrence, all births, deaths, marriages, and migrations, beginning in June 1974 and continuing to today. HDSS included censuses in 1974 and 1982 that provided evidence of preprogram similarities and differences between treatment and comparison areas. It also allowed us to identify women present in the study area at program rollout in October 1977. To reduce potential bias due to selective in-migration into treatment versus comparison areas and to avoid issues related to endogeneity, we use location prior to rollout to assign treatment status. The full inclusion table (*SI Appendix*, Fig. S2) shows that 59,189 women born 1938–1973 met these criteria.

Second, MHSS1, conducted in 1996, included a 7% sample of Matlab HDSS households. At the time of MHSS1, 32,869 eligible women (55.5%) remained in the HDSS area, with exits largely due to migration of women at marriage. MHSS1 surveyed 2,104 eligible women. Using HDSS data, we tested for differences between treatment and comparison areas in mortality or in exit from the HDSS between October 1977 and start of MHSS1 in April 1996 and in probability of selection into MHSS1, and none were statistically significant.

Finally, the principal data source was MHSS2, a panel follow-up of all individuals surveyed in MHSS1. The sample was tracked in HDSS through December 2014, with subsequent intensive in-person follow-up and migrant tracking, which enabled detailed documentation of mortality timing and yielded very low attrition rates. Of 2,104 eligible MHSS1 respondents, 1,820 completed MHSS2 interviews (86.5% of all MHSS1 respondents, 94.8% of surviving MHSS1 respondents). The combined MHSS1/MHSS2 analysis file is available for download. There were no significant treatment−comparison differences in mortality for 1996–2012 for any cohort (*SI Appendix*, Table S2).

Measures.

Fertility outcomes. Measures of completed family size, age at first birth, birth spacing, age at last birth, and contraceptive use were included to analyze lifetime programmatic impacts on fertility and on key pathways of effects on fertility and to test their roles as potential mechanisms explaining health impacts. Timing of first birth and last birth, average birth interval, children ever born, and surviving children were constructed from MHSS2 pregnancy histories. These histories were initially prepopulated with prospectively observed data from HDSS, augmented with births occurring outside Matlab recorded in MHSS2, and cross-checked against births reported in MHSS1. Ever use of any modern contraception, injectable contraceptive, intrauterine device (IUD), oral contraceptive pill, and sterilization were constructed from contraceptive life histories in MHSS1 and MHSS2.

Health and economic outcomes. We constructed normalized indices for three domains—metabolic health, functional health, and respiratory health—and summed them to provide an index of overall health status. Higher values of each index indicate poorer health. In addition to analyzing each index, we conducted analyses of its component variables. We also considered measures that were available in both MHSS1 and MHSS2, including BMI and overweight and underweight status, to connect our findings to earlier MHSS1 results and to better understand MHSS2 findings. Economic outcomes include household consumption, women's individual earnings, employment status, and savings. Construction of health measures is described in *Materials and Methods*, with detail on all variables in *SI Appendix*, sections 4 and 7.

Independent variables. Treatment status is an ITT indicator based on respondent's area of residence (treatment or comparison) prior to rollout. We used demographic and socioeconomic measures from the 1974 and 1982 censuses to establish baseline balance and to serve as control variables. Details on key variable construction are provided in *SI Appendix*, section 4.

Estimation Strategy. We estimated ITT effects of the MCH-FP program on outcomes in MHSS2 for women who were of reproductive age during the first decade of the program. We estimated single-difference models on all women and on the three birth cohorts chosen by age at first exposure to MCH/FP. We used ordinary least squares regression, with SEs clustered at the preprogram village level to account for the likely intracluster correlation in the error terms. All models controlled for baseline characteristics, most of which were balanced in 1974 but are included to account for possible preprogram sociodemographic and exposure differentials, as described in *SI Appendix*, Table S1. Tubewell access was not balanced at baseline; we therefore also controlled for tubewell arsenic level using icddr,b 2003 measures to further account for the confounding effects of widespread arsenic contamination (14, 15) All models were adjusted for attrition between October 1977 and MHSS2 using inverse probability weights estimated using baseline variables, birth year, river embankment, flood risk, and treatment status (*SI Appendix*, section 9). To check robustness of our results, we tested models that accounted for additional key confounders, that varied the cohort groupings, and that used alternate weighting schemes, as described in *Materials and Methods*.

Effects on Fertility and Contraceptive Use. We begin by describing ITT effects of the MCH/FP program on key measures of lifetime fertility and contraceptive uptake. Table 1 presents treatment and comparison area means for the three cohorts separately, regression-adjusted differences in means, and CIs for the adjusted difference. We found both significant treatment−comparison differences and a downward cohort trend in measures of family size. Adjusted for covariates, women in MCH/FP treatment area versus comparison area had 0.52 fewer births in the oldest cohort (unadjusted mean T = 6.66, C = 7.20, 95% CI for adjusted difference −0.94 to −0.10, $P < 0.05$), 0.67 fewer births in the middle cohort (unadjusted mean T = 4.95, C = 5.72, 95% CI −0.98 to −0.35, $P < 0.01$), and 0.50 fewer in the youngest cohort (unadjusted mean T = 3.51, C = 4.30, 95% CI −0.81 to −0.20, $P < 0.01$). Differences between children born and children surviving were considerably smaller for the youngest two cohorts, due to successful child mortality interventions introduced first in the treatment area.

The effects of the program on birth spacing evolved over time. We observed no significant area differences in age at first birth, but the youngest cohort began childbearing later. Age at last birth declined across cohorts, and women in the treatment area ceased childbearing at significantly younger ages in both the 1938–1949 cohort (T = 34.61, C = 36.02, adjusted diff = −1.42, 95% CI −2.35 to −0.48, $P < 0.01$) and the 1950–1961 cohort (T = 31.63, C = 32.88, adjusted diff = −1.00, 95% CI −1.95 to −0.04, $P < 0.05$). Treatment area women had significantly longer average birth intervals in the 1950–1961 cohort (T = 3.32, C = 3.02, adjusted diff = 0.30, 95% CI 0.14 to 0.47, $P < 0.01$) and the 1962–1973 cohort (T = 4.44, C = 3.56, adjusted diff = 0.81, 95% CI 0.49 to 1.13, $P < 0.01$), but not in the oldest cohort.

Table 1 also captures the large increase in contraceptive use over time and the substantial early impact of the program on use, with a cohort pattern of emerging treatment area advantage followed by narrowing differences as family planning arrived in the comparison area.

Effects on Later-Life Health. Fig. 1 summarizes MCH/FP treatment effects from regression models for the four health indices for all women and for the three cohorts. Poorer health is indicated by positive effects. Looking first at all ages combined, overall health was significantly poorer in the treatment area by +0.07 SDs (95% CI 0.016 to 0.136, $P < 0.05$), which was driven by a +0.12 SD difference in the respiratory domain (95% CI 0.010 to 0.226, $P < 0.05$) and a +0.05 SD difference in the metabolic domain (95% CI −0.017 to 0.122, $P = 0.138$). When we examined each birth cohort separately, a modest pattern of poorer health among the treated emerged for the middle cohort. In this cohort, born 1950–1961, there was a +0.11 SD (95% CI 0.013 to 0.201, $P < 0.05$) difference in the overall health domain and a +0.15 SD difference (95% CI 0.048 to 0.255, $P < 0.01$) in the metabolic domain. The effects are unpacked within each domain in *SI Appendix*, section 5; they show a consistent pattern of treatment area disadvantage in the middle cohort across multiple metabolic disease outcomes (angina, diabetes, etc.).

To understand the unfolding of metabolic consequences of MCH/FP over time, Table 2 presents effects of MCH/FP on BMI and underweight (<18.5 kg/m^2), overweight/obese (Asian standard, >23 kg/m^2), and obese (Asian standard, >27.5 kg/m^2) indicators across survey round and across cohort. In MHSS1, each cohort/treatment group had a mean BMI of 19.6 kg/m^2 or under and 41–66% were underweight after adjusting for baseline covariates. Treatment area BMI was significantly higher only in the middle age cohort (0.76 kg/m^2, 95% CI 0.29 to 1.22, $P < 0.01$). This weight gain was associated with nine percentage points lower underweight status (95% CI −0.17 to 0.00, $P < 0.10$) and six percentage points higher overweight status (95% CI 0.01 to 0.10, $P < 0.05$). In the oldest cohort, treatment area women were 13 percentage points lower in underweight status (95% CI −0.22 to −0.03, $P < 0.01$). No other weight differences were significant.

Table 2 also presents 2012 MHSS2 results. Average BMI had increased to a range of 20.3 kg/m^2 to 22.5 kg/m^2 across the three cohorts, and underweight was reduced by nearly half or more, to 15 to 33%. Overweight had emerged for all women, increasing from a range of 2 to 12% in MHSS1 to a range of 23 to 43% in MHSS2. For the 1950–1961 cohort, the MCH/FP effect on BMI was positive but only borderline significant (0.57 kg/m^2, 95% CI −0.04 to 1.18, $P < 0.10$) and the distributional effects of that weight gain had shifted. Treatment area women were seven percentage points more likely to be overweight/obese (95% CI 0.00 to 0.15, $P < 0.10$). While marginally significant, the increase in overweight/obese was nonetheless one component of the significant difference in metabolic index shown in Fig. 1. Effects of MCH/FP on underweight, overweight/obese, and obese were not significant for other groups. We analyzed the mediating factors explaining associations between MCH/FP and later health outcomes (*SI Appendix*, section 6), and found that controlling for BMI in 1996 explained about one-third of the total metabolic effect.

Finally, we analyzed associations between MCH/FP exposure and economic outcomes at both MHSS1 and MHSS2. After adjusting for covariates, we found few significant effects of MCH/FP on consumption, employment, earnings, or assets for the combined group or for the individual cohorts (*SI Appendix*, section 7).

Discussion

This study explored the effects, up to 35 y later, of the highly influential Matlab MCH/FP program on multiple dimensions of health and economic well-being among women who were childbearing age at or within 10 y of the program's inception in 1977. The program accomplished its primary objective of enabling improved fertility control: We confirm earlier observed highly significant differences in completed fertility and lifetime contraceptive behavior overall and when dividing the population into three cohorts born 1938–1949, 1950–1961, and 1962–1973. A related study showed that the program was associated with increased human capital and skilled employment among children born during the program period (16, 17). Yet, surprisingly, we found few long-term program effects on women's health or economic outcomes, with the exception of a consistent and significant association with poorer metabolic and functional health among women born 1950–1961. We discuss failure to detect benefits separately from negative impact.

The limited long-term benefit of the program may reflect the fact that the Matlab program primarily aimed to avert higher-order births among current mothers in a population with early

SOCIAL SCIENCES

ECONOMIC SCIENCES

Table 1. Program effects on fertility and contraceptive behavior: Differences between treatment and comparison area means by birth cohort

	1938–1949				1950–1961				1962–1973			
	Mean		Adjusted diff		Mean		Adjusted diff		Mean		Adjusted diff	
	C	T	T-C	N	C	T	T-C	N	C	T	T-C	N
Family size												
Children born	**7.20**	**6.66**	−0.52*	552	**5.72**	**4.95**	−0.67**	697	**4.30**	**3.51**	−0.50**	570
			[−0.94, −0.10]				[−0.98, −0.35]				[−0.81, −0.20]	
Surviving children	5.41	5.24	−0.24	552	**4.73**	**4.22**	−0.40**	697	**3.71**	**3.15**	−0.36**	570
			[−0.55, 0.06]				[−0.66, −0.14]				[−0.63, −0.09]	
Birth spacing												
Age at first child	18.39	18.77	0.23	543	18.81	18.73	−0.09	685	20.86	21.55	0.37	559
			[−0.43, 0.89]				[−0.53, 0.35]				[−0.34, 1.07]	
Age at last child	**36.02**	**34.61**	−1.42**	543	**32.88**	**31.63**	−1.00*	685	32.25	31.85	−0.08	559
			[−2.35, −0.48]				[−1.95, −0.04]				[−1.01, 0.86]	
Average birth interval	2.99	2.94	−0.02	537	**3.02**	**3.32**	0.30**	676	**3.56**	**4.44**	0.81**	542
			[−0.24, 0.20]				[0.14, 0.47]				[0.49, 1.13]	
Ever use of modern contraception												
Any	**0.15**	**0.48**	0.31**	552	**0.57**	**0.83**	0.26**	698	**0.58**	**0.77**	0.19**	570
			[0.23, 0.40]				[0.19, 0.33]				[0.11, 0.28]	
Injection	**0.03**	**0.32**	0.28**	552	**0.22**	**0.70**	0.48**	698	**0.47**	**0.73**	0.26**	570
			[0.21, 0.35]				[0.41, 0.54]				[0.19, 0.34]	
IUD	**0.01**	**0.05**	0.04+	552	**0.07**	**0.16**	0.08*	698	**0.05**	**0.10**	0.04*	570
			[−0.00, 0.08]				[0.02, 0.14]				[0.00, 0.09]	
Pill	**0.11**	**0.21**	0.07*	552	**0.43**	**0.55**	0.13**	698	0.76	0.75	0.02	570
			[0.00, 0.14]				[0.04, 0.21]				[−0.07, 0.10]	
Sterilization	**0.08**	**0.17**	0.06*	552	**0.21**	**0.15**	−0.06*	698	0.16	0.12	−0.05	570
			[0.00, 0.13]				[−0.12, −0.01]				[−0.11, 0.01]	

C, comparison group mean; T, treatment group mean. Unadjusted means in bold differ between areas at $P < 0.05$. Adjusted diff denotes model-adjusted difference between treatment and comparison areas, adjusted for religion, 1974 baseline characteristics (own age, household head's education, household head's spouse's education, bari size, family size, dummies for household head's occupation type [agriculture, fishing, business], dummies for household characteristics [tin walls, tin roof, latrine, tubewell water, tank water, number of rooms per capita], assets [number of boats, number of cows, household owns a lamp, owns a watch, owns a radio]) and land sized owned in 1982, and categorical controls for 2003 arsenic exposure at 50 ppb to 99 ppb, 100 ppb to 149 ppb, 150 ppb to 399 ppb, and >400 ppb; n, number of observations in group. $+P < 0.10$; $*P < 0.05$; $**P < 0.01$.

onset of childbearing. The intervention may have been too late in the reproductive life course for these women. The few existing studies observing long-term nonfertility benefits of family planning were focused on interventions to delay the onset of childbearing and prevent later unplanned pregnancies. None of the cohorts in this study experienced differentials in the onset of childbearing, although there was a secular change toward delayed childbearing in the youngest cohort born 1962–1973. This cohort did have significantly longer birth intervals in the treatment area, however, and still did not experience differential long-term health effects.

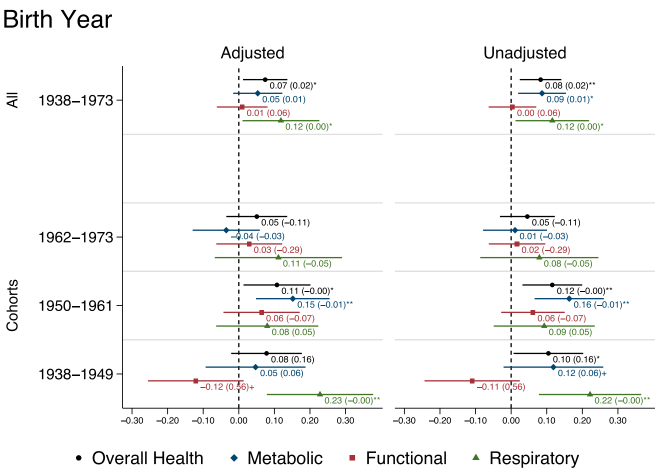

Fig. 1. Program effects on health domains: Differences between treatment and comparison area means for all women born 1938–1973 and by birth cohort. $+P < 0.10$, $*P < 0.05$, $**P < 0.01$. Error bars represent 95% confidence intervals.

The finding of modestly significant negative health consequences among women reaching peak childbearing age during the program was surprising. These negative health consequences can perhaps be understood as the persistence of the BMI effects of MCH/FP (4) under a rapidly shifting nutritional transition. This transition, which could not have been anticipated in 1977, became a growing cause of concern through much of South Asia over the course of the study period. In 1996, mean BMI for women in the treatment area was 19.6 kg/m^2 versus 18.7 kg/m^2 in the comparison area. As a result, a large percentage of women in the two oldest cohorts were lifted out of underweight status (Table 2). The nutritional transition is illustrated in Fig. 2, which tracks the BMI distribution in our middle cohort at two time points. In the approximately 16 y between MHSS1 and MHSS2, there was a massive increase in mean BMI of over 2.5 kg/m^2 in both treatment and comparison areas. Mean BMI at the later timepoint was 22.2 kg/m^2 in the treatment area compared to 21.4 kg/m^2 in the comparison area. The treatment-comparison difference seen in MHSS1 remained in MHSS2, although smaller. The BMI distribution in MHSS1 was highly kurtotic with little variation; in MHSS2 it both shifted right and showed dramatically increased dispersion. As a result, a BMI effect that had once served to reduce underweight (in MHSS1) was associated with increased

Table 2. Program effects on BMI and weight: Differences between treatment and comparison area means by birth cohort

	1938–1949				1950–1961				1962–1973			
	Mean		Adjusted diff		Mean		Adjusted diff		Mean		Adjusted diff	
	C	T	T-C	N	C	T	T-C	N	C	T	T-C	N
MHSS1												
BMI	**18.17**	**18.88**	0.45	472	**18.66**	**19.63**	0.76**	618	18.82	19.04	0.13	492
			[−0.23, 1.13]				[0.29, 1.22]				[−0.31, 0.58]	
Underweight	**0.66**	**0.52**	−0.13**	472	**0.53**	**0.41**	−0.09+	617	0.45	0.42	−0.04	490
			[−0.22, −0.03]				[−0.17, 0.00]				[−0.14, 0.06]	
Overweight or obese	0.06	0.08	−0.01	472	**0.04**	**0.12**	0.06*	617	0.02	0.05	0.02	490
			[−0.06, 0.04]				[0.01, 0.10]				[−0.02, 0.05]	
Obese	0.01	0.01	0.00	472	0.01	0.02	0.01	617	0.00	0.01	0.00	490
			[−0.03, 0.03]				[−0.01, 0.03]				[−0.02, 0.02]	
MHSS2												
BMI	20.27	20.84	0.27	524	**21.41**	**22.22**	0.57+	679	21.97	22.50	0.28	547
			[−0.43, 0.97]				[−0.04, 1.18]				[−0.38, 0.94]	
Underweight	0.33	0.29	−0.03	524	0.23	0.19	−0.03	679	0.19	0.15	−0.04	547
			[−0.12, 0.05]				[−0.10, 0.04]				[−0.11, 0.02]	
Overweight or obese	0.23	0.26	0.01	524	**0.32**	**0.42**	0.07+	679	**0.34**	**0.43**	0.04	547
			[−0.07, 0.09]				[−0.00, 0.15]				[−0.05, 0.13]	
Obese	0.03	0.05	0.00	524	0.07	0.10	0.01	679	0.10	0.10	−0.01	547
			[−0.03, 0.04]				[−0.03, 0.06]				[−0.06, 0.05]	

See Table 1 legend. MHSS1 entries report MHSS1 outcomes of those who responded to MHSS2.

overweight (in MHSS2). It was accompanied by a small excess in metabolic risk and self-reported angina in MHSS2.

The lack of economic differentials is also somewhat surprising. An earlier study based on MHSS1 data found that wages for employed women were higher in 1996 among women who lived in the treatment area, but rates of employment were lower (18). Our analysis is based on the ITT design described above and uses total earnings across all economic activities. We find no effects on economic outcomes, as shown in *SI Appendix*, section 7. Other MHSS1/MHSS2 studies of men in these cohorts also find no significant effects on their employment and earnings (17). This does not mean we can fully rule out past economic differentials. The 16-y gap between study rounds means that economic effects with persistent consequences, including for BMI, could have peaked and fallen before 1996 or between study rounds.

At the same time, it is possible that the BMI effects resulted from other pathways such as access to health services that treated infections and diarrheal diseases, thus reducing associated weight loss. Also possible are direct metabolic effects of injectable contraceptive products (19) or effects of changes in women's time use and activities due to changing family composition.

This study had other limitations. First, assignment to MCH/FP was block level instead of randomized, in order to avoid contamination effects. Although the study areas were shown to be highly similar at baseline and we controlled for a number of historic confounders, it remains possible that some unobserved factor suppressed long-term effects in the treatment area. Second, the use of a wedge design in which services were extended to the comparison area after roughly a dozen years eliminates the possibility of a fully untreated comparison group, although this limitation would exist in any long-term study of family planning, given the importance of these interventions and the long duration of women's reproductive years. Third, the study did not include biomarker-based measures of cardiovascular disease or diabetes. Finally, since the nutritional transition was universal in this population, we cannot possibly know how this evaluation would have played out in the absence of such change.

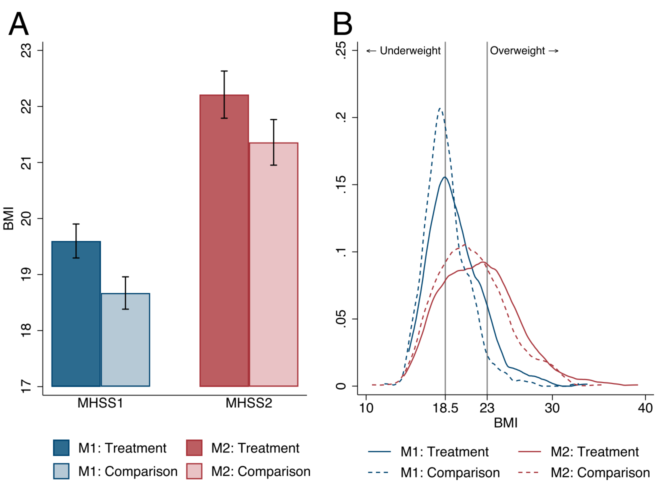

Fig. 2. Means (*A*) and distributions (*B*) for BMI of 1950–1961 MHSS2 cohort at MHSS1 and MHSS2. Error bars represent 95% confidence intervals.

Nevertheless, our findings from a long-term quasi-random intervention should lead us to be cautious when anticipating long-term benefits of family planning programs, and to be prepared both for large-scale societal changes such as nutritional transition to swamp potential long-term benefits or even have negative unintended consequences. Family planning programs are a universal good unto themselves because they avert unwanted pregnancies, reduce maternal mortality, and accord reproductive freedom. Yet the institutional, cultural, and epidemiological impediments to long-term indirect effects are quite profound, and childbearing also accords benefits to women in many societies. Even where the MCH/FP program yielded benefits in the form of increased BMI and reduced underweight, it may also have hastened the arrival of metabolic risk.

Since these interventions were introduced more than 40 y ago, it would be easy to question their external validity to the modern family planning landscape. Yet a good share of today's global family planning agenda looks remarkably similar to the Matlab MCH/FP program. While many programs target adolescents with a broad range of sexual health and livelihood options, the primary focus still rests with large-scale interventions targeting married women with services aimed at fertility stopping behavior. While this approach may be necessary and strategic, our findings suggest that we should not place unreasonably high expectations on programs focused on stopping behavior. The primary goals of such narrowly focused programs are to reduce family size and reduce short-term exposure to excess maternal mortality risk, and that may be all we can or should expect. Family planning programs may be more likely to achieve additional long-term benefits if they are embedded in a system directly focused on women's life course health and well-being.

Materials and Methods

Measurement. The three health indices included the following items derived from MHSS2: metabolic—indicators for angina and stroke (based on self-reported symptoms), overweight or obesity (based on the Asian standard, BMI $\geq$ 23.0 kg/m^2 measured using scale and Shorr board) and stage 1 and stage 2 hypertension (based on standard blood pressure cutoffs) (20); functional—indicators for self-reported problems with ADLs and poor health status and measures using objective physical functioning tests and dynamometers for poor physical capability score (<13), adult daily living mobility score (0 to 18), and maximum dominant hand grip strength (kilograms); and respiratory—indicators for chronic lung disease and asthma (based on self-reported symptoms).

We constructed each index by standardizing each component variable using the comparison group mean and SD. We then took the arithmetic mean of the standardized component values. The sign of a component variable was switched as necessary so that higher values represent worse health outcomes. Additional information on the measures, including cutoffs or thresholds used to construct the components, is provided in *SI Appendix*, section 4.

Robustness Checks. We also accounted for several other important and well-documented changes that occurred in Matlab over the 35-y period since program inception that could confound results, including the introduction of a microcredit experiment in half of the treatment and comparison areas (21, 22), construction of a river embankment that altered cropping patterns and flood risks in protected areas (23), and the gradual urbanization of villages near Matlab town. These factors had no substantial impact on results and so were not included in final models. We also examined models with no 1974 controls and models with controls for women's migration behavior; results were similar to our included models. Details on the confounders we considered and the results of these models are in *SI Appendix*, section 8.

To check the robustness of estimates, we also tested models using different cohort groupings (see *SI Appendix*, section 8). As noted above, exposure to MCH/FP treatment was assigned based on village of residence in 1974, prior to program implementation. To replicate earlier work (4), we also estimated models that assigned treatment based on area of residence in 1996 or 2012, and these results did not differ substantially. We also estimated models that removed the attrition weights. Additional detail on weight construction and the results of this robustness check are provided in *SI Appendix*, section 9, showing that results were similar with no weighting.

Data Availability. The analysis file for this study is available for download and reanalysis at openICPSR (https://www.openicpsr.org/openicpsr/project/143101/version/V1/view) (24).

ACKNOWLEDGMENTS. This study was funded and supported by the US National Institute on Aging, the International Initiative for Impact Evaluation, the Population Reference Bureau, and the Eunice Kennedy Shriver National Institute of Child Health and Human Development. We thank icddr,b for collaboration and access to Matlab HDSS data, including vital registration data from 1974–2014; the 1974 and 1982 censuses; and GIS and arsenic test data. We thank Nobuko Mizoguchi, Sveta Milusheva, Michael Emch, and Omar Rahman for their contributions to the design and collection of MHSS2 data. We thank the staff of Mitra and Associates for their leadership in collecting MHSS1 and MHSS2, particularly S. N. Mitra, Shahidul Islam, Fuad Pasha, Manir Bhuiyan, Zahidul Islam, and Syed Abdullah Al Ahsan. We thank icddr,b colleagues and other advisory board members who provided input into study and questionnaire design, including Sajeda Amin, Abbas Bhuiya, Zafrullah Chowdhury, John Clemens, Shams El Arifeen, Terry Hull, Ruchira Tabassum Naved, Nurul Alam, Anisur Rahman, Mizanur Rahman, Rubhana Raqib, Laura Reichenbach, Kim Streatfield, and the late Simeen Mahmud. We gratefully acknowledge icddr,b staff who were instrumental in completing the survey, including Sajal Kumar Saha, Taslim Ali, Jharna Akter, Kamal Hossain, Monoranjan Das, A. H. M. Golam Mustafa, and, especially, the dear departed Nazrul Islam. We also thank the University of Colorado Boulder support team, especially Steve Graham. Our design benefited from historical insights on the formation of the MCH/FP program from Muhammad Yunus, S. Chakraborty, A. M. Sardar, Jim Phillips, and Henry Mosley. Earlier versions of the paper were presented at International Union for the Scientific Study of Population, Population Association of America, International Conference on Family Planning, and Asian Population Association meetings. Most of all, we thank the residents of the Matlab HDSS area and those who answered surveys even after leaving their area for their nearly 60-y commitment to global health science.

1. D. Canning, T. P. Schultz, The economic consequences of reproductive health and family planning. *Lancet* **380**, 165–171 (2012).
2. J. E. Finlay, M. A. Lee, Identifying causal effects of reproductive health improvements on women's economic empowerment through the Population Poverty Research Initiative. *Milbank Q.* **96**, 300–322 (2018).
3. Z. Zimmer, J. Kwong, Family size and support of older adults in urban and rural China: Current effects and future implications. *Demography* **40**, 23–44 (2003).
4. S. Joshi, T. P. Schultz, Family planning and women's and children's health: Long-term consequences of an outreach program in Matlab, Bangladesh. *Demography* **50**, 149–180 (2013).
5. T. Barham, Enhancing cognitive functioning: Medium-term effects of a health and family planning program in Matlab. *Am. Econ. J. Appl. Econ.* **4**, 245–273 (2012).
6. J. F. Phillips, R. Simmons, M. A. Koenig, J. Chakraborty, Determinants of reproductive change in a traditional society: Evidence from Matlab, Bangladesh. *Stud. Fam. Plann.* **19**, 313–334 (1988).
7. J. Menken, J. F. Phillips, Population change in a rural area of Bangladesh, 1967-87. *Ann. Am. Acad. Pol. Soc. Sci.* **510**, 87–101 (1990).
8. M. A. Koenig, U. Rob, M. A. Khan, J. Chakraborty, V. Fauveau, Contraceptive use in Matlab, Bangladesh in 1990: Levels, trends, and explanations. *Stud. Fam. Plann.* **23**, 352–364 (1992).
9. M. A. Koenig, V. Fauveau, A. I. Chowdhury, J. Chakraborty, M. A. Khan, Maternal mortality in Matlab, Bangladesh: 1976-85. *Stud. Fam. Plann.* **19**, 69–80 (1988).
10. G. Miller, Contraception as development? New evidence from family planning in Colombia. *Econ. J. (Lond.)* **120**, 709–736 (2010).
11. V. Fauveau, "Matlab: physical setting and cultural background" in *Matlab: Women, Children and Health, ICDDR, B Special Publication*, V. Fauveau, Ed. (Pioneer, 1994), pp. 13–28.
12. D. H. Huber, A. R. Khan, Contraceptive distribution in Bangladesh villages: The initial impact. *Stud. Fam. Plann.* **10**, 246–253 (1979).
13. W. S. Stinson, J. F. Phillips, M. Rahman, J. Chakraborty, The demographic impact of the contraceptive distribution project in Matlab, Bangladesh. *Stud. Fam. Plann.* **13**, 141–148 (1982).
14. M. Rahman *et al*., Prevalence of arsenic exposure and skin lesions. A population based survey in Matlab, Bangladesh. *J. Epidemiol. Community Health* **60**, 242–248 (2006).
15. M. Yunus, N. Sohel, S. K. Hore, M. Rahman, Arsenic exposure and adverse health effects: A review of recent findings from arsenic and health studies in Matlab, Bangladesh. *Kaohsiung J. Med. Sci.* **27**, 371–376 (2011).
16. T. Barham, R. Kuhn, Staying for benefits: The effect of a health and family planning program on out-migration patterns in Bangladesh. *J. Hum. Resour.* **49**, 982–1013 (2014).
17. T. Barham, R. Kuhn, J. Menken, A. Razzaque, "Thirty-five years later: Evaluating the impacts of a child health and family planning programme in Bangladesh" (Impact Evaluation Rep. 39, International Initiative for Impact Evaluation, New Delhi, India, 2016).
18. T. P. Schultz, How does family planning promote development? Evidence from a social experiment in Matlab, Bangladesh, 1977-1996. http://www.econ.yale.edu/~pschultz/TPS_10_30_QJE.pdf. Accessed 22 June 2021.
19. A. B. Berenson, M. Rahman, Changes in weight, total fat, percent body fat, and central-to-peripheral fat ratio associated with injectable and oral contraceptive use. *Am. J. Obstet. Gynecol.* **200**, 329.e1–329.e8 (2009).
20. Mayo Clinic, High blood pressure (hypertension) - Diagnosis and treatment. https://www.mayoclinic.org/diseases-conditions/high-blood-pressure/diagnosis-treatment/drc-20373417. Accessed 22 September 2020.
21. S. Masud Ahmed, M. Chowdhury, A. Bhuiya, Micro-credit and emotional well-being: Experience of poor rural women from Matlab, Bangladesh. *World Dev.* **29**, 1957–1966 (2001).
22. A. Bhuiya, M. Chowdhury, Beneficial effects of a woman-focused development programme on child survival: Evidence from rural Bangladesh. *Soc. Sci. Med.* **55**, 1553–1560 (2002).
23. A. M. Mobarak, R. Kuhn, C. Peters, Consanguinity and other marriage market effects of a wealth shock in Bangladesh. *Demography* **50**, 1845–1871 (2013).
24. T. Barham *et al*., Data and code for: Thirty-five years later: Long-term effects of the Matlab maternal and child health/family planning program on older women's well-being. openICPRS. https://www.openicpsr.org/openicpsr/project/143101/version/V1/view. Deposited 16 June 2021.