Title

Comparative effectiveness of two disparate policies on child health: experimental evidence from the Philippines.

Journal

Health Policy and Planning, 32(4)

Authors

Peabody, John
Quimbo, Stella
Florentino, Jhiedon
et al.

Publication Date

2017-05-01

DOI

10.1093/heapol/czw179

Peer reviewed

Health Policy and Planning, 32, 2017, 563–571
doi: 10.1093/heapol/czw179
Advance Access Publication Date: 20 January 2017
Original Article

OXFORD

Comparative effectiveness of two disparate policies on child health: experimental evidence from the Philippines

John W Peabody,[1,2,*] Stella Quimbo,[3] Jhiedon Florentino,[2] Riti Shimkhada,[4] Xylee Javier,[3] David Paculdo,[2] Dean Jamison[5] and Orville Solon[3]

[1]Global Health Sciences, University of California, San Francisco, CA, USA, [2]QURE Healthcare, San Francisco, CA, USA, [3]University of the Philippines, School of Economics, Phillippines, [4]Center for Health Policy Research, University of California, Los Angeles, CA, USA, [5]Department of Global Health, University of Washington, WA, USA

*Corresponding author. John W. Peabody, Global Health Sciences, University of California, San Francisco, and QURE Healthcare, 450 Pacific St. Suite 200, San Francisco, CA 94133, USA. E-mail: jpeabody@qurehealthcare.com

Accepted on 19 December 2016

Abstract

Background Should health systems invest more in access to care by expanding insurance coverage or in health care services including improving the quality of care? Comparing these options experimentally would shed light on the impact and cost-effectiveness of these strategies.

Methods The Quality Improvement Demonstration Study (QIDS) was a randomized policy experiment conducted across 30 districts in the Philippines. The study had a control group and two policy intervention groups intended to improve the health of young children. The demand-side intervention in QIDS was universal health insurance coverage (UHC) for children aged 5 years or younger, and a supply-side intervention, a pay-for-performance (P4P) bonus for all providers who met predetermined quality levels. In this paper, we compare the impacts of these policies from the QIDS experiment on childhood wasting by calculating DALYs averted per US$spent.

Results The direct per capita costs to implement UHC and P4P are US$4.08 and US$1.98 higher, respectively, compared to control. DALYs due to wasting were reduced by 334,862 in UHC and 1,073,185 in P4P. When adjustments are made for the efficiency of higher quality, the DALYS averted per US$ spent is similar in the two arms, 1.56 and 1.58 for UHC and P4P, respectively. Since the P4P quality improvements touches all patients seen by qualifying providers (32% in UHC versus 100% in P4P), there is a larger reduction in DALYs. With similar programmatic costs for either intervention, in this study, each US$spent under P4P yielded 1.52 DALYs averted compared to the standard program, while UHC yielded only a 0.50 DALY reduction.

Conclusion P4P had a greater impact and was more cost-effective compared to UHC as measured by DALYs averted. While expanded insurance benefit ceilings affected only those who are covered, P4P incentivizes practice quality improvement regardless of whether children are insured or uninsured.

Keywords: Comparative effectiveness, pay for performance, Philippines, policy experiment, universal health coverage

Introduction

Among the most basic policy debates in health is whether to intervene on the demand side, supporting patients or consumers, or to intervene on the supply side, supporting providers. Economists and practitioners often see the 'patient versus provider' investment debate—which approach spends how much on which group—staged in ideological or intuition terms rather than on the merits or research

Key Messages

- In an experimental setting, this study examined the cost effectiveness of two policies intended to improve the health of young children - a demand-side intervention introducing universal health insurance coverage versus a supply-side pay for performance intervention to improve quality of care – compared to controls.
- Comparing costs over Disability Adjusted Life Years (DALYs), this study found the supply-side pay for performance intervention had a greater impact and was more cost-effective compared to the demand-side universal health insurance intervention.
- As policy makers are forced to choose between seemingly good ideas, this study suggests that comparative effectiveness research, typically limited to medial/clinical sciences, could be expanded to comparing differences in larger system level health policy options.

evidence. Despite the significance of these fundamentally different choices, there is sparse evidence comparing supply versus demand policies (Gopalan *et al.* 2014, Lundberg *et al.* 2006, Musgrove 2011, Wharam and Daniels 2007).

Wasting, defined as low weight for height, is an important population health indicator used for monitoring nutritional status and health in child populations globally. Children suffer from wasting and growth retardation because of malnutrition resulting from poor diets, more frequent episodes of severe diarrhoea and susceptibility to infectious diseases, such as pneumonia (Rodriguez *et al.* 2011). Severe malnutrition is a significant contributor to child mortality globally (Bhutta 2009). There is strong evidence that impaired growth is also associated with the morbidities of delayed mental development, poor school performance, and reduced intellectual capacity (Kar *et al.* 2008, Laus *et al.* 2011). Wasting affects about 52 million children globally (Black *et al.* 2013). In the Philippines, 960 900 children were wasted in 2010 (Food and Nutrition Research Institute 2008), and in the Central Philippines wasting affected approximately 7% of all children 5 years and under.

Between 2003 and 2009, we conducted a randomized controlled policy experiment in the Central Philippines, widely known as QIDS (the Quality Improvement Demonstration Study), that compared a demand-side intervention—universal health coverage (UHC)—with a supply-side intervention—a pay-for-performance (P4P) scheme for physicians caring for children under 6 years of age. Previous reports from the QIDS study showed both the UHC and P4P interventions improved wasting over time compared to the control group (Peabody *et al.* 2014, Quimbo *et al.* 2011). In this analysis, we build upon these findings and use a DALYs per US$spent framework to measure the direct impact of the two policies on childhood wasting and compare the cost of each policy. This paper thus presents a comparison of the intervention costs and the benefits on DALYs due to wasting to determine which intervention provided the biggest impact for the smallest cost.

Methods

Data collection

The QIDS social policy experiment (ClinicalTrials.Gov (#NCT00678197)) involved a partnership between the Philippine Health Insurance Corporation (PhilHealth), the Department of Health and an academic team from the University of the Philippines and the University of California San Francisco (Shimkhada *et al.* 2008). The QIDS study was conducted in accordance with the ethical standards of the applicable national and institutional review boards (IRBs) of the University of the Philippines and the University of California, San Francisco (CHR Approval Number: H10609-19947-05). QIDS implementation and evaluation, funded by the U.S. National Institutes of Health as the Philippine Child Health Experiment (NICHD #R01HD042117) was conducted in 30 communities and corresponding district hospitals located in 4 different regions in the Central Philippines (the Visayas and Mindanao islands). The study experimentally introduced into the 30 communities either 1) UHC for children under 5 years, 2) quality-based P4P incentives for physicians caring for children under 5 years old and younger in district-level hospitals or 3) no policy change (the control or standard policy sites). Both QIDS policy interventions were financed and operationalized by PhilHealth, the government funded organization responsible for the national health insurance program.

We collected baseline data on child health beginning in 2003. Post-intervention data were collected in 2007, three years after the two policies were introduced. Voluntary participation was very high, and informed consent was secured from all parents of the participants. Overall in both rounds of data collection, 89% of the parents agreed to participate. We administered to the participants' parents an in-hospital survey at discharge, 4–10 weeks post-discharge, and another in the follow-up in-home survey. Trained interviewers obtained a wide range of information including socio-demographic information, diagnosis, treatment and health status. We measured health status by obtaining anthropometric data to concurrently assess for wasting. Following the work of others in the Philippines and other countries, we defined wasting as having an actual weight of child to be $< 90\%$ of their ideal weight for height (Del Mundo 1999). We measured height using stiff measuring tapes, lying children down (those under 2 years) or standing them up (those over 2 years). We measured weight using 25 kg Salter scales designed for field use, or with standing hospital scales in patient exit surveys. We measured height and weight twice for each child.

At baseline, there were 1,011 patients in the control arm, 985 in the UHC sites, and 992 in the P4P sites; in the second survey round, there were 1,031 children in the control arm, 1,042 in the UHC sites, and 1,036 in the P4P sites.

Policy interventions

The first policy (known as UHC or the demand side intervention) targeted children 5 years old and younger by providing expanded health insurance coverage. 'Policy Navigators', individuals who were tasked to ensure that the intervention was effectively implemented, ensured that enrollment was universal. Specifically, we tasked them to enroll indigent households into PhilHealth by liaising with local town mayors, provincial governors and other officials (Solon *et al.* 2009). Moreover, insurance benefit ceilings were raised for children 5 years old and younger when admitted to the district hospitals in the UHC study sites. Using our baseline data on hospital

Table 1. Inputs used to calculate YLL, YLD, and DALYs due to wasting among children under 5 years of age, applied to total Philippines population

		Values
Number of children under 5[a]	A	10,231,201
0–11 months	A1	1,967,576
12–59 months	A2	8,263,625
Years of Life Lost		
Mortality rates (NSCB, Philippine MDG Goals) (per thousand)		
Infants (0–11 months)	B1	31.0
Under 5 (12–59 months)	B2	23.0
Number of deaths in children under 5	C = (A1*B1) + (A2+B2)	251,058
Global estimate of deaths due to wasting for children under 5 (%)[21]	D	14.60%
Number of deaths in children under 5 years due to wasting	E=C*D	36,654
Standard years of life remaining at age of death (Calculated as average standard life of 68.96 years subtracted by five years due to the average duration of malnutrition until remission)	F	63.96 years
YLL due to wasting	**G=E*F**	**2,344,422**
Years Living with Disability		
Prevalence of wasting among children under 5 (%)[b]	H	26.26%
Number of under 5 wasted	I=A*H	2,686,713
Estimated probability that a child under 5 years suffering from wasting will die	J	1.36%
Average duration of malnutrition until remission or death	K	5 years
Disability weight for wasting[c]	L	0.053
YLD due to wasting	**M=I*J*K*L**	**702,296**
DALYs due to wasting (sum of YLL and YLD due to wasting)	**N=G+M**	**3,046,688**

[a]2010 Census of Population and Housing. Philippine Statistical Authority.
[b]Baseline wasting for a random control site (HH).
[c]Global Burden of Disease 2004 Update: Disability Weights for Diseases and Conditions. WHO.

prices, with the increase in insurance benefit ceilings, we estimated that young children enjoyed full insurance coverage in any UHC study site.

The second policy (known as P4P or the supply side intervention) introduced a bonus scheme for doctors and their staff. Bonus eligibility was pre-defined for the doctors using a three-component metric: the quality of the clinical care (70%), patient satisfaction (20%) and a minimum workload (number of patients seen per week) (10%). The bonus opportunity represented approximately a 5% emolument relative to total physician salaries.

We measured the quality of clinical care using Clinical Performance and Value (CPV®) vignettes. CPVs are an established, affordable and well-validated measure of clinical performance that has been used globally (Peabody *et al.* 2014, Peabody et al. 2004, 2008, 2011). Individual CPV scores of the quality of care, combined with patient satisfaction and workload (*i.e.* the P4P metric) were detailed and fed back to the doctor and the hospital director. Aggregated scores for each district hospital were given to the hospital director and to the provincial public health officials, including the governors, introducing transparency around performance. To increase overall sustainability and mitigate policy costs, no specific interventions, such as training or education, were offered in the P4P intervention sites. The expectation was that the incentive and the transparency would be adequate to improve health outcomes compared to controls.

Determination of cost estimates

To calculate the per claim costs of the two interventions, we used the QIDS project financial statements and PhilHealth records from the 2014 PhilHealth Annual Report and PhilHealth records on QIDS-identified claims and reimbursements. The total cost of UHC includes the marginal cost of insurance, the administrative costs of PhilHealth (which we refer to as 'direct costs'), the QURE program administration cost for control and intervention sites, and the additional program administration cost of the Policy Navigators used in UHC. The total cost of P4P includes the cost of insurance claims, the administrative costs of PhilHealth the QURE program administration cost and the bonus payments to those facilities that qualified for the performance incentive. We obtained the standard program administration costs and value of insurance claims filed in all project sites from the official PhilHealth records. In the case of P4P intervention sites, PhilHealth records also provided the costs of the incentive payments to qualifying participating hospitals. The additional intervention administration costs such as the employment of three Policy Navigators at UHC sites and the expenses associated with generating the performance quality measures (*e.g.* rostering, interview of physicians with vignette administration, scoring of vignettes, encoding, dissemination of results to hospital chiefs, governors and mayors, staff costs including analysis) were obtained from QIDS project records.

We determined the unit costs of each intervention by dividing the total cost of each intervention by the number of program beneficiaries. To assess the unit cost of the interventions, we compare these with the unit cost of the PhilHealth program in the control sites.

Standard Program. The standard administration costs were assumed to be 6.5% of the total value of insurance claims, which is estimated from the Statement of Profit of Loss of PhilHealth in 2014. The total value of claims is based on the observed average value of claim from the end line patient exit surveys of QIDS.

Universal Health Coverage. UHC had the additional program administration costs of employing Policy Navigators to ensure

Table 2. Cost calculation (in US$) of rolling out policies at a national level

		Control	UHC	P4P
Direct Cost (Value of Insurance Claims)				
Population, under 5 years old[1]	A	10,231,201		
% covered by PhilHealth among under 5 years old	a	57.81%		
Number of children, under 5 years old, covered by PhilHealth	A1=A*a	5,914,657		
% who were sick and confined, under 5 years old[b]	B	7.64%		
Number of children, under 5 years old, who were sick and got confined	C=A*B	781,664		
Of sick and confined children, under 5 years old, % who PhilHealth covered and claimed[c]	D	25.64%	31.57%	37.15%
Number of children, under 5 years old, sick, confined, PhilHealth beneficiaries, and claimed	E=C*D	200,419	246,771	290,388
Average value of PhilHealth claims (in US$)*,[d]	F	45.39	49.15	57.76
Decrease in charges due to quality improvements[e]	q			40.00%
Decrease in the value of insurance claims due to quality improvement (in US$)	f = (1-q)*F			23.11
Average value of insurance claim (with quality adjustment) (in US$)	G=F-f	45.39	49.15	34.66
Cost of Program Administration				
Program administration cost (in US$)[f]	H1=G*E*6.5%	$591,305	$788,372	$654,140
Costs of Navigators in UHC sites (US$0.8 per household, 0.16 per person in a household[13] (in US$))[g]	H2=A1*0.16		$946,345	
Costs of bonus payments P4P sites[h] (~US$2.04 per patient, in US$)	H3=E*US$2			$592,392[h]
Total cost of program administration (in US$)	H = (E*G*6.5%) + G1+G2	$591,305[f]	$1,734,718[g]	$1,246,532
Unit cost (per program beneficiary)				
Program administration cost	J1=H1/E	$2.95	$3.19	$2.25
Navigator cost in UHC	J2=H2/E		$3.84	
Bonus payments in P4P	J3=H3/E			$2.04
Total per unit cost of rolling out the intervention (in US$)	J=J1+J2+J3	$2.95	$7.03	$4.29 ($5.79 if w/o quality adjustment)

*We converted expenditures to 2015 USD using a ratio of PhP46/US$1.

[a]Source: Census of Population, 2010.

[b]Source: 2013 National Demographic and Health Survey.

[c]Source: QIDS Patient Exit Survey, endline (2006).

[d]Source: QIDS Patient Exit Survey, endline (2006); with adjustment from inflation using consumer prices indices from the Philippine Statistical Authority website (http://www.nscb.gov.ph/secstat/d_price.asp).

[e]Peabody et al. Quality Variation and its Impact on Costs and Satisfaction: Evidence from the QIDS Study. Med Care 2010; 48 25–30.

[f]The control sites are assumed to have a program administration cost of 6.5%, which is the share of program administration costs in total benefit payments, according to the PhilHealth Annual Report 2014. Great Leaps: Charting the Future of Philippine Health Care (www.philhealth.gov.ph).

[g]Assuming 5 household members, the cost of navigation is US$0.16 per individual.

[h]Bonus payments to doctors are estimated by multiplying the bonus rate per patient day and length of stay, and the assumed proportion of doctors qualifying for bonuses (80%, which was observed during the last round of QIDS).

enrollment of families into the insurance program, which amounted to about 35 pesos (US$0.8) per household covered and was included in the cost of the intervention (Solon *et al.* 2009). In a household of five members (average), this amounts to US$0.16 per member in a covered household.

Pay for Performance. In the P4P sites, the additional program administration costs were for the bonuses. For costing purposes, we used an 80% passing rate equal to the quality threshold accomplished at the end of the QIDS study. Additionally, general practitioners belonging to facilities that passed the quality measure received a daily US$2.17 increase in the quality-care fee. Based on the QIDS Exit Patient Survey, at an average length of stay of three days, this amounts to US$6.52 per patient in quality-care professional fees. At the last round of the quality assessment survey, 80% of the facilities passed. Applying these proportions in the estimated US$6.52 above, qualified facilities under P4P gives an average of US$2.04 per patient as a bonus.

Consistent with previous findings that doctors providing higher quality tend to be more prudent in prescribing drugs and avoid ordering unnecessary laboratory tests, we accounted for this in the course of the QIDS study. We found that the total charges of patients, who were cared for by QIDS physicians with higher quality scores, was 40% lower compared to patients cared for by physicians with lower quality scores, and we adjusted the costs per patient in the P4P sites accordingly (James *et al.* 2009).

Calculation of YLL, YLD, DALY

We used the 2010 data on the infant mortality of the Philippines population to estimate the deaths and DALYs averted among those aged under 5. To estimate the impact of wasting on deaths, we relied on the global estimate of the share of wasting on deaths for children under 5 of 14.6%,[21] and the estimated probability a child under 5 years suffering from wasting who will die is 1.36%. To calculate the

Table 3. Intervention impacts on wasting and impacts on YLL, YLD, DALYs

	UHC	P4P
Intervention-specific parameters		
Utilization rate for the intervention	32.00%	100%
Number of wasted children in the intervention groups[a]	2,686,713	
Estimated probability that an under 5 suffering from wasting will die	1.36%	
Absolute reduction in wasting due to intervention	9.0%[b]	9.25%[c]
Percent wasting given the intervention	17.26%	17.01%
Number of wasted children, post-intervention	2,356,687	1,716,584
Reduction of wasting due to intervention	293,372	933,475
Averted deaths due to wasting	4,020	12,907
Summary		
YLL due to wasting	2,087,303	1,518,607
YLD due to wasting	624,522	454,895
DALYs	2,711,825	1,973,502
Reduction in DALYs (because of wasting) due to Intervention	334,862	1,073,186

[a]Baseline wasting for a random control site (HH).
[b]Significant at 5% (Martins et al. 2011).
[c]Significant at 5% (Peabody et al. 2000).

Table 4. Computation of cost effectiveness ratios

		Control	UHC	P4P
DALYs	A	3,046,688	2,711,825	1,973,502
Total cost of program administration (in million US$)	B (H in 3)	0.59	1.73	1.25
Total per unit cost of rolling out the intervention (in US$)	C (in 3)	Cc = 2.95	Cuhc=7.03	Cp4p=4.29
DALY per US$spent in program administration	D=A/B	Dc = 5.15	Duhc=1.56	Dp4p=1.58
Percentage change in DALY per US$relative to control	E = (Di-Dc)/Dc(i=uhc,p4p)		−69.67	−69.28
Percentage change in unit cost relative to control	F = (Ci-Cc)/Cc(i=uhc,p4p)		138.29	45.50
Decrease in DALY per US$for every increase in unit cost (relative to control)	G=E/F		−0.50	−1.52

potential years of life lost (YLL) prior to the intervention, the number of deaths is multiplied by the standard life expectancy at age of death in years. The average standard life expectancy in the Philippines is 68.96 (Central Intelligence Agency 2013) which decreases by five years due to the long-term effects of malnutrition (Schroeder and Brown 1994, Martins 2011).

From this five-year effect of malnutrition, we also assumed that the average years lived with disability (YLD) due to malnutrition is five years. This assumption is likely conservative, as it assumes no excess YLD over the decrease in life expectancy. Calculating the YLD is thus the product of the number of cases, disability weight for wasting, and the average duration of the case until remission or death. The disability weight of 0.053 is from the WHO Global of Burden and Disease, 2004 update. To avoid double counting in the DALY computation, the estimated YLD does not include the number of children who will die due to wasting.

For this study, we calculated the estimated DALYs averted from wasting due to each intervention among children under 5 years old. DALYs are a combined measure of the overall mortality and morbidity burden brought about by a specific disease or condition, namely (a) the potential YLL, and (b) the YLD. For the interventions, the reduction of DALYs under UHC is dependent on whether the sick children are insured, confined, and utilize their insurance coverage, while under P4P the reduction in DALYs is dependent only on whether the sick children are confined.

Intervention impact on reduction of DALYs

We projected the population level estimates by applying the findings of the QIDS interventions onto the population data. The 2010 population census data on the number of children (aged 0 up to 5 years old) were used. The number of children suffering from wasting was calculated using prevalence data from the QIDS baseline random household survey. The total number of deaths among infants (0 to 11 months) and children (12 to 59 months) were calculated based on the Millennium Development Goal (MDG) results on infant and child mortality. Using the global estimate of 14.6% of the share of wasting on deaths for children under 5 deaths (Black *et al.* 2008), the number of deaths due to wasting was estimated at 1.36% of the number of children suffering from wasting.

As noted previously, the overall QIDS design was a randomized controlled trial of UHC versus P4P involving 30 districts in the Visayas region of the Philippines (Peabody *et al.* 2014). Previous studies with these data substantiated that the QIDS interventions produced health effects on its targeted children (Peabody *et al.* 2014, Quimbo *et al.* 2011). UHC was found to be linked to a 9% to 12% reduction in wasting and a 4% to 9% reduction of having an infection (p-value < 0.001 for both) (Quimbo *et al.* 2011). P4P produced a similar 9.25% reduction in wasting (p-value < 0001), as well as a 7% increase in subjective health rating (p-value = 0.001), relative to the control sites.[11] These findings and the temporal concurrence of the two policy interventions are used to enumerate DALYs averted in the Philippine population from wasting due to the experimentally applied interventions among children <5 years old.

Cost-effectiveness analysis

Cost-effectiveness analysis was determined from the perspective of PhilHealth, or the administrator of the policy programs. That is, we wanted to determine which policy—whether one of the

Table 5. Analysis of intervention effectiveness under different coverage and quality assumptions

	% covered and claimed	DALYs per US$spent	DALYs averted per US$spent
UHC	26%	1.74	0.40
	29%	1.67	0.45
	32%	1.56	0.50
	% of facilities passed	DALYs per US$spent	DALYs averted per US$spent
P4P	70%	1.84	1.41
	85%	1.71	1.47
	100%	1.58	1.52

interventions or the standard program—was the most effective at reducing the effects of wasting for each US$spent. To do this, we needed to calculate the DALYs due to wasting under the standard policy, subtract the number of DALYs averted due to each intervention, then divide by the total cost of implementing either the standard policy or the intervention. Secondarily, for each intervention group, we examined the percentage change in DALYs per US$spent relative to the standard policy and divided by the percentage change in per unit cost relative to the standard policy. The resulting value is the number of DALYs averted for each US$spent.

Results

Calculating DALYs

The total population for children under 5 was 10 231 201 (1). The total number of deaths among infants (0 to 11 months) and children (12 to 59 months) were calculated based on the MDG results on infant and child mortality and the total population of children under 5 years old, yielding 251 058 total deaths in this population. Using the global estimate deaths due to wasting for children under 5 years old, 14.6%, the number of deaths due to wasting was estimated to be 36 654 children. Calculating the standard years of life remaining at age of death to be 63.96 years, this means there is a total of 2,344 422 YLL due to wasting.

Based on the wasting prevalence in a random control site in the QIDS study at baseline, the prevalence of wasting among children under 5 was assumed to be 26.26%, which means 2,686 713 children under 5 years were suffering from wasting. To avoid double counting, we used an estimated probability of 1.36% that a child under 5 years suffering from wasting will die. Assuming an average duration of malnutrition until remission or death of 5 years and using a disability weight of 0.053 for wasting, this translates to 702 266 YLD due to wasting.

Using the numbers generated, we arrive at a baseline of 3,046 688 DALYs due to wasting under the standard program.

Intervention impacts on DALYs

As shown in 1, without either intervention, 36 654 of the children nationwide would die because of wasting. This is equivalent to 2,344 422 YLLs. Our previous work found that an absolute reduction in wasting prevalence from the baseline of 26.26% could be achieved by both interventions and by nearly the same amount: 9.0% in UHC sites and 9.25% in P4P sites. Under UHC, 32% of the population was covered by health insurance intervention, resulting in 4,020 fewer deaths due to wasting. Assuming the same 63.96 years of standard life remaining, this is equivalent to 2,087 303 YLLs due to wasting.

Under P4P, all patients of P4P qualifying physicians are impacted and thus all patients receive better care (100% utilization rate for the intervention), and 12 907 deaths due to wasting are averted, reducing the total deaths due to wasting to 23 743. This is equivalent to 1,518 607 YLLs due to wasting.

There is a similarly important policy impact on YLDs from wasting. The total YLDs without either intervention is 702 266, but is reduced to 624 522 and 454 895 for UHC and P4P, respectively. This means that the number of wasted children is reduced from 2.687 million to 2.357 and 1.717 million under UHC and P4P respectively, if applied to the entire population under 5.

The reduction in DALYs is 334 862 under UHC and 1,073 186 under P4P, resulting in 2,711 825 and 1,973 502 residual DALYs for UHC and P4P, respectively. Understanding that the gains in health due to UHC apply only to those who use their coverage to obtain care, the lower number of averted DALYs due to UHC is based on the utilization rate of 32% garnered from the QIDS Exit Patient survey. This drives the lower value for UHC compared with P4P. We estimated the level of utilization needed for UHC sites to achieve the same level of impact in DALYs as P4P sites, not until insurance utilization reaches 100% does the reduction in DALYs converge for UHC and P4P.

The estimated health effects roughly translate to averting 2,880 and 9,250 cases of wasting for every 100 thousand children under 5 in UHC sites versus P4P sites, respectively. Assuming that the interventions become a national policy, this reduction in wasting translates to 293 372 and 933 475 fewer wasted children, for UHC and P4P respectively. Additionally, the rate per thousand of children dying due to wasting goes from 3.58 under the standard policy to 3.19 for UHC and 2.32 for P4P.

Policy cost estimates

2 shows the estimated costs of rolling out the intervention policies at the national level, composed of the administration cost related to the processing of claims (standard program administration cost for all three programs), the cost of ensuring enrollment to the program (UHC only), and the cost of bonus payments for doctors (P4P only).

As of the 2010 census, there were 10.23 million children under 5, of whom 7.64% became sick and were confined to hospital. From the QIDS Exit Patient Survey, we know that the percentage of eligible participants claimed by PhilHealth and the average value of the claim for each policy were 25.64% and US$45.39, 31.57% and US$45.19, and 37.15% and US$57.76 under standard program, UHC and P4P, respectively. Utilizing the 40% decrease in charges due to quality improvements under P4P, reduces the average value of insurance claims for P4P to US$34.66.

Administration costs for each program is the average value of insurance claim (with quality adjustment) multiplied by the number of children under 5 years who were sick, confined, covered and claimed by PhilHealth (2, row H1) multiplied by 6.5% (the assumed standard program administration costs). Factoring in the cost of Policy Navigators in UHC (found by multiplying US$0.16 times the percent of under 5 years old covered by PhilHealth) and

the cost of bonus payments in P4P (found by multiplying US$2.04 by the total number of claims for this population) yields the total cost of program administration for the two intervention policies. Thus, on a per unit cost to implement each program (2, rows J2 and J3), P4P (US$2.04) is lower than UHC (US$3.84) by 47%.

Cost-effectiveness

Using the cost calculations in 2, on a per unit cost basis, the standard control program (US$2.95) will be lower compared to both UHC (US$7.05) and P4P (US$4.29).

To measure the value of the programs, including impact on health burden, we used the computed the YLLs, YLDs, and DALYs averted due to wasting per US$spent for children under 5. As inputs, we used the percentage of deaths due to wasting and the number of years suffering from malnutrition from 1. For the standard program, this yielded a baseline DALYs of 3,046 688 (1). We then calculated the reduction in DALYs attributed to each intervention and determined that the residual DALYs are 2,711 825 for UHC and 1,973 502 for P4P. That is, UHC reduced the baseline DALYs by 334 862 and P4P reduced DALYs by 1,073 186 (3). Using the baseline and residual DALYs, we next determined the DALYs due to wasting per US$spent (4, Row D). This shows that the highest ratio for the standard program (control group) at 5.15 DALYs per US$, substantially higher than either UHC (1.56 DALYs per US$) or P4P (1.58 DALYs per US$). In other words, while the standard program is much lower in direct costs, the disease burden is much higher than in UHC or P4P as measured in absolute DALYs or DALYs per US$basis(4).

We note that both interventions are nearly equivalent in terms of DALYs per dollar spent (1.56 and 1.58 for UHC and P4P, respectively). To differentiate between the interventions, we evaluate which policy results in a greater decrease in DALYs per incremental increase in US$spent. We first computed the percentage decrease in DALYs per US$for each intervention, in comparison to the standard program. These were nearly equivalent for the interventions: 69.7% for UHC and 69.3% for P4P (4). We then compared the per unit (direct) costs for each intervention and find that UHC costs are 138.29% higher and P4P costs are 45.50% higher versus the standard program. Taking the ratio of decrease in DALYs to increase in US$for each intervention yields −0.50 for UHS and −1.52 for P4P. This means every one US$spent implementing UHC returned a reduction of 0.50 DALYs. By contrast, everyone US$spent in implementing P4P led to an associated reduction of 1.52 DALYs.

Sensitivity analysis

We performed a sensitivity analysis on DALYs per US$spent and DALYs averted per US$spent, under two additional scenarios for each policy (5). We know from our analysis above that UHC impact is affected by the proportion of children who are covered and able to claim their benefits, while the effectiveness of P4P is affected by the proportion of doctors and facilities passing a quality threshold. For UHC, which was 26% coverage under the standard policy, we raised the number of patients claiming benefits to 29% (versus 32% in the actual study). Similarly, for P4P we raised the number of facilities that passed to 85% (versus 100% in the actual study) where 70% is the percentage of providers who passed under the standard policy. We find that as coverage increases, the DALYs averted per US$spent increases. In the case of UHC, this increase in DALYs averted per US$spent comes from reducing the fixed cost of the Policy Navigators, which is distributed among more patients. Under P4P, the increase in DALYs averted per US$spent is due to a greater number of patients receiving higher quality care from qualifying facilities. At 100% coverage (and utilization of their coverage) the DALYs averted under UHC is equivalent to the DALYs averted under P4P albeit at much higher cost.

Discussion

Translating promising research findings, particularly those that advance health or lower health care costs, into scale interventions is a core challenge in health today. These important stakes—and the paucity of translational successes—have led to increasing calls for evidence-based policymaking (Brownson *et al.* 2009, Jacobs *et al.* 2012). Ideally, evidence-based policy making is comparative, evaluating real-world economic efficiencies that allows a determination of policy value and not just policy efficacy (Teutsch and Fielding 2011).

Concurrently, there are vigorous, new interests in determining comparative cost effectiveness of clinical interventions using comparative effectiveness research. Comparative effectiveness studies, however, typically focus only on therapeutic interventions and not policy approaches. In this study, we were able to compare the effectiveness of a demand policy intervention with a supply side policy intervention. We found that both interventions were effective and both interventions—relative to controls—reduced wasting by about 9%. Interestingly the costs in the supply-side intervention were notably less than the demand side intervention. These results argue that in the setting of constrained resources, increasing quality is more cost effective than expanding insurance benefits at improving health status in this population of children. This finding is supported by a body of literature on the effectiveness of supply side externalities, which likely operate through standardization of clinical care, more appropriate testing and decreased time in the hospital (Bodenheimer and Fernandez 2005).

When benefits and costs were first viewed in isolation, the QIDS results initially suggested that both policies were equally effective and had similar costs. However, when we do a cost-effectiveness analysis (herein), two important factors change these estimates decidedly in favour of a policy to increase quality using financial incentives. First, when quality externalities are taken into account, there is a much greater impact from quality manifested by a greater reduction in disease burden DALYs and thus a potentially large impact on reducing child mortality. Our analysis also shows, too, that cost-effectiveness differences disappear as universal coverage is approached. Thus, when benefits are similar, the costs of universal coverage alone determine the relative cost effectiveness. Second, while the costs of the two policies appear to be similar on a cost per patient, supply-side P4P has the added benefit of improving quality and lowering the cost of care for all patients cared for by P4P clinicians when UHC is incomplete. The impact of the demand-side UHC intervention is limited just to those patients that would otherwise not have accessed care.

There are several caveats to interpreting these findings. First, cost structures are inherently hard to ascertain and could be structurally different in other settings, for example if premiums were not actuarially set and greatly exceeded expected case costs. We do not believe that is the case here, as premiums are intentionally less than the price of goods and services. Costs, in general, were based upon billed charges not actual costs, and thus are almost certainly understate the true inputs. However, this underestimation is unlikely to differentially affect the two policies. Second, despite the health benefits of the performance incentives and to a lesser degree insurance

expansion, the overall benefits of these two policies are underestimated relative to other low and middle income countries where the percentage of children who suffer from severe caloric deficiency is higher. By only looking at wasting, our estimates further err on the conservative because we did not take into account milder cases of under-nutrition, known to affect 130 million additional children worldwide (Bhutta 2009). A potentially more appropriate measure, for future studies, is stunting rather than wasting. Stunting captures chronic undernutrition, but requires a longer study to be able to capture impacts. The policy interventions might also affect other diseases as manifest by other outcomes that we failed to capture. And, certainly there may be other benefits that we did not determine, such as other co-morbidities. Our motivation to use a measure of malnutrition is based upon its ubiquity across preven/trea diseases. We attempted to look at differences in mortality, but even with the larger sample size there was not enough power to generate a meaningful analysis. Further research would be useful to determine if the interventions affected disease incidence, related costs such as length of stay, readmissions and rehabilitation.

Because wasting often turns into longer term undernutrition that leads to stunting and impaired growth, wasted children are also likely to have other long term benefits such as decreased chronic disease later in life, reduced allostatic burden and higher IQ effects from better nutrition and better school attendance (Kar et al. 2008, Ampaabeng and Tan 2013). There are other externalities that we did not measure such as sibling effects from reduced disease burden, indirect cost savings from parents that can return to work, better nutrition, and long term gains in human capital (Jamison 2006, Miguel and Kremer 2004). We were not able to calculate, for example, increases in non-health consumption. We expect, however, that these externalities would benefit the entire family providing greater financial risk protection and better quality care.

We did not perform a full enumeration and costing of these externalities as it goes beyond the scope of this study and exceeds the breadth of our data. However, we feel it is quite likely that the calculated benefits derived herein are underestimated, particularly for access to quality care. From other studies, for example, we know that patients are willing to travel farther to obtain quality healthcare (Leonard 2014). By introducing a quality scheme throughout an entire region, travel costs would be reduced. In this way, the overall benefit derived from P4P could be even more pronounced than is found in this study.

As policy makers are forced to choose between seemingly good ideas, this study suggests that comparative effectiveness research should be expanded to looking at differences in policy options. This study compared two disparate policies and found clear overall support for a supply side intervention not often considered in resource poor settings. This finding may be useful to make policy choices and allocate precious healthcare resources.

Conclusion

Using experimental data, this policy evaluation study shows the comparative costs and effects of two very different types of interventions, one using the demand approach and the other using the supply approach. The results suggest both the demand-side and the supply-side intervention improve childhood wasting. However, it is the supply side intervention, the P4P strategy, which appears to be the most cost effective.

Ethics committee approval

Study conducted in accordance with the ethical standards of the applicable national and institutional review boards (IRBs) of the University of the Philippines and the University of California, San Francisco (CHR Approval Number: H10609-19947-05)

Funding

The Quality Improvement Demonstration Study was funded by the U.S. National Institutes for Health through an R01 grant (HD042117). ClinicalTrials.Gov (#NCT00678197)

Conflict of interest statement. This work was conducted through the University of California, San Francisco under a U.S. National Institute of Health grant and with a subcontract to the University of the Philippines. After the conclusion of this study and writing of this report, author John W. Peabody founded QURE Healthcare. QURE Healthcare is a privately owned company (President, John W. Peabody) that uses the performance measure described in this report, the Clinical Performance and Value (CPV®) vignettes, in a proprietary manner, in different clinical settings than described in this study.

References

Ampaabeng S, Tan C. 2013. The long-term cognitive consequences of early childhood malnutrition: the case of famine in Ghana. *Journal of Health Economics* **32**: 1013–27.

Bhutta Z. 2009. Addressing severe acute malnutrition where it matters. *Lancet* **374**: 94–6.

Black RE, Allen LH, Bhutta ZA *et al*. 2008. Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* **371**: 243–60.

Black RE, Victora CG, Walker SP *et al*. 2013. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* **382**: 427–51.

Bodenheimer T, Fernandez A. 2005. High and rising health care costs. Part 4: Can costs be controlled while preserving quality? *Annals of Internal Medicine* **143**: 26–31.

Brownson R, Fielding J, Maylahn C. 2009. Evidence-based public health: a fundamental concept for public health practice. *Annual Review of Public Health* **30**: 175–201.

Central Intelligence Agency. The World Factbook 2013-14. Washington, DC: Central Intelligence Agency, 2013. https://www.cia.gov/library/publications/the-world-factbook/fields/2102.html. Accessed January 7, 2016.

Del Mundo F. *Textbook of Pediatrics and Child Health*. Vol 4th ed. Quezon City, Philippines: JMC Press; 1999.

Food and Nutrition Research Institute, (Philippines) D of S and T. Philippines National Nutrition Survey 2008-2009.; 2008.

Ginsburg PB. 2004. Controlling Health Care Costs. *The New England Journal of Medicine* **351**: 1591–3.

Gopalan S, Mutasa R, Friedman J, Das A. 2014. Health sector demand-side financial incentives in low- and middle-income countries: a systematic review on demand- and supply-side effects. *Socical Science & Medicine* **100**: 72–83.

Jacobs J, Jones E, Gabella B, Spring B, Brownson R. 2012. Tools for implementing an evidence-based approach in public health practice. *Preventing Chronic Disease* **9**: E116.

James CD, Peabody J, Solon O, Quimbo S, Hanson K. 2009. An unhealthy public-private tension: pharmacy ownership, prescribing, and spending in the Philippines. *Health Affairs (Milliwood)* **28**: 1022–33.

Jamison DT, Investing in Health. In: Jamison DT, Breman JG, Measham AR *et al*. eds. *Disease Control Priorities in Developing Countries*. Vol 2nd ed. Washington, DC; 2006.

Kar B, Rao S, Chandramouli B. 2008. Cognitive development in children with chronic protein energy malnutrition. *Behavioral Brain Function* **4**: 31.

Laus M, Vales L, Costa T, Almeida S. 2011. Early postnatal protein-calorie malnutrition and cognition: a review of human and animal studies. *International Journal of Environmental Research and Public Health* **8**: 590–612.

Leonard KL. 2014. Active patients in rural African health care: implications for research and policy. *Health Policy Plan* **29**: 85–95.

Lundberg M, Wang L. Health Sector Reforms. In: Coudouel A, Paternostro S, eds. *Analyzing the Distributional Effects of Reform*. Vol Washington, DC: World Bank; 2006:43–100.

Martins VJB, Toledo Florê TMM, Grillo LP *et al*. 2011. Long-lasting effects of undernutrition. *International Journal of Environmental Research and Public Health* **8**: 1817–46. doi:10.3390/ijerph8061817.

Miguel E, Kremer M. 2004. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica* **72**: 159–217.

Musgrove P. *Financial and Other Rewards for Good Performance or Results: A Guided Tour of Concepts and Terms and a Short Glossary*. Washington, DC: World Bank; 2011.

Peabody JW, Shimkhada R, Quimbo S *et al*. 2014. The impact of performance incentives on child health outcomes: results from a cluster randomized controlled trial in the Philippines. *Health Policy Plan* **29**: 615–21. doi:10.1093/heapol/czt047.

Quimbo SA, Peabody JW, Shimkhada R, Florentino J, Solon O. 2011 May. Evidence of a causal link between health outcomes, insurance coverage, and a policy to expand access: experimental data from children in the Philippines. *Health Economics* **20**: 620–30. doi: 10.1002/hec.1621.

Rodriguez L, Cervantes E, Ortiz R. 2011. Malnutrition and gastrointestinal and respiratory infections in children: a public health problem. *Journal of Environmental Research and Public Health* **8**: 1174–205.

Shimkhada R, Peabody J, Quimbo S, Solon O. 2008. The quality improvement demonstration study: an example of evidence-based policy-making in practice. *Health Research Policy and Systems* **6**: 5.

Solon O, Peabody JW, Woo K *et al*. 2009. An evaluation of the cost-effectiveness of policy navigators to improve access to care for the poor in the Philippines. *Health Policy (New York)* **92**: 89–95.

Wharam J, Daniels N. 2007. Toward evidence-based policy making and standardized assessment of health policy reform. *JAMA* **298**: 676–9.

Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. 2000. Comparison of vignettes, standardized patients, and chart abstraction: A prospective validation study of 3 methods for measuring quality. *JAMA* **283**: 1715–22. doi:10.1001/jama.283.13.1715.

Peabody J, Shimkhada R, Quimbo S *et al*. 2011. Financial incentives and measurement improved physicians' quality of care in the Philippines. *Health Affairs* **30**: 773–81. doi:10.1377/hlthaff.2009.0782.

Peabody JW, Luck J, Glassman P *et al*. 2004. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Annals of Internal Medicine* **141**: 771–80.

PhilHealth. Great Leaps: Charting the Future of Philippine Health Care. Annual Report 2014. http://www.philhealth.gov.ph/about_us/annual_report/ar2014.pdf.

Schroeder DG, Brown KH. 1994. Nutritional status as a predictor of child survival: summarizing the association and quantifying its global impact. *Bulletin of the World Health Organization* **72**: 569–75.

Teutsch S, Fielding J. 2011. Applying comparative effectiveness research to public and population health initiatives. *Health Affairs* **30**: 349–55.