NBER WORKING PAPER SERIES

ARE QUALITY-ADJUSTED MEDICAL PRICES DECLINING FOR CHRONIC DISEASE? EVIDENCE FROM DIABETES CARE IN FOUR HEALTH SYSTEMS

Karen Eggleston
Brian K. Chen
Chih-Hung Chen
Ying Isabel Chen
Talitha Feenstra
Toshiaki Iizuka
Janet Tinkei Lam
Gabriel M. Leung
Jui-fen Rachel Lu
Beatriz Rodriguez-Sanchez
Jeroen N. Struijs
Jianchao Quan
Joseph P. Newhouse

WORKING PAPER 25971

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2019

We are grateful to the individuals and data providers for allowing access to the individual-level data and to our comparative health system team coauthors for the individual system-level analyses for ability to undertake this comparative study. Karen Eggleston gratefully acknowledges support from a Stanford FSI Shorenstein Asia-Pacific Research Center faculty research award; for the Japan study, Toshiaki Iizuka, Brian Chen and Karen Eggleston gratefully acknowledge funding the Abe Fellowship Program, JSPS Bilateral Open Partnership Joint Research Projects, and a Shorenstein APARC faculty research award; for the Netherlands study, Beatriz Rodriguez-Sanchez, Talitha Feenstra, and Jeroen N. Struijs gratefully acknowledge Professor H. Bilo for allowing access to the ZODIAC data and Chantal van Tilburg and Mirte van Galen from VEKTIS for providing hospitality to work on the linked dataset; for the Hong Kong study, Jianchao Quan, Janet Tinkei Lam, and Gabriel Leung gratefully acknowledge funding from Research Grants Council of the Hong Kong SAR, China (Project No. 27112518) and thank the Hospital Authority for providing data; and for the Taiwan study, Juifen Rachel Lu, Ying Isabel Chen, Chih-Hung Chen, Brian Chen and Karen Eggleston gratefully acknowledge Chang Gung Memorial Hospital for providing the patient-level data for analyses and grant funding support from the Taiwan Ministry of Science and Technology (MOST 100- 2632-H-182-001-MY2, 104-2918-I-182-002). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Are Quality-Adjusted Medical Prices Declining for Chronic Disease? Evidence from Diabetes Care in Four Health Systems
Karen Eggleston, Brian K. Chen, Chih-Hung Chen, Ying Isabel Chen, Talitha Feenstra, Toshiaki Iizuka, Janet Tinkei Lam, Gabriel M. Leung, Jui-fen Rachel Lu, Beatriz Rodriguez-Sanchez, Jeroen N. Struijs, Jianchao Quan, and Joseph P. Newhouse
NBER Working Paper No. 25971
June 2019
JEL No. H51,I10,I18

ABSTRACT

Improvements in medical treatment have contributed to rising health spending. Yet there is relatively little evidence on whether the spending increase is “worth it” in the sense of producing better health outcomes of commensurate value—a critical question for understanding productivity in the health sector and, as that sector grows, for deriving an accurate quality-adjusted price index for an entire economy. We analyze individual-level panel data on medical spending and health outcomes for 123,548 patients with type 2 diabetes in four health systems. Using a “cost-of-living” method that measures value based on improved survival, we find a positive net value of diabetes care: the value of improved survival outweighs the added costs of care in each of the four health systems. This finding is robust to accounting for selective survival, end-of-life spending, and a range of values for a life-year or, equivalently, to attributing only a fraction of survival improvements to medical care.

Karen Eggleston
Shorenstein Asia-Pacific Research Center
Stanford University
616 Serra Street
Stanford, CA 94305
and NBER
karene@stanford.edu

Brian K. Chen
Department of Health Services Policy
and Management
Arnold School of Public Health
University of South Carolina
915 Greene Street #354
Columbia, SC 29205
bchen@mailbox.sc.edu

Chih-Hung Chen
Department of Endocrinology
and Metabolism
Chang Gung Memorial Hospital
Keelung
yh1008@cgmh.org.tw

Ying Isabel Chen
Graduate Institute of Epidemiology
National Taiwan University
Taiwan
glamorous2238@gmail.com

Talitha Feenstra
National Institute for Public Health and
Environment (RIVM)
Hague
The Netherlands
talitha.feenstra@rivm.nl

Toshiaki Iizuka
Graduate School of Economics
University of Tokyo
Bunkyo-ku, Japan
toshi.iizuka@gmail.com

Janet Tinkei Lam
University of Hong Kong
Hong Kong SAR
janetltk@connect.hku.hk

Gabriel M. Leung
School of Public Health
University of Hong Kong
Hong Kong SAR
gmleung@hku.hk

Jui-fen Rachel Lu
Department of Health Care Management
Chang Gung University
Taoyuan City, Taiwan
rachel@mail.cgu.edu.tw

Beatriz Rodriguez-Sanchez
Universidad de Castilla-La Mancha
Spain
Beatriz.RSanchez@uclm.es

Jeroen N. Struijs
National Institute for Public Health
and Environment
Leiden University
Medical Center Campus
The Hague
The Netherlands
jeroen.struijs@rivm.nl

Jianchao Quan
University of Hong Kong
Hong Kong SAR
jquan@hku.hk

Joseph P. Newhouse
Division of Health Policy Research
and Education
Harvard University
180 Longwood Avenue
Boston, MA 02115-5899
and NBER
newhouse@hcp.med.harvard.edu

A data appendix is available at http://www.nber.org/data-appendix/w25971

Medical spending has increased significantly in most economies in recent decades, challenging health system financing and prompting efforts to control cost growth. Improvements in medical treatment have clearly contributed to significant increases in medical spending, yet there is relatively little quantitative evidence on whether the rise in expenditure is "worth it" in the sense of producing better health outcomes of commensurate value. In the absence of such evidence, simply imposing cost controls may stifle life-saving innovations as much as reduce low-value care and "waste."

As a result, understanding and enhancing the net value of medical spending is of substantial importance for assessing and improving productivity in the growing service sectors of most economies. A few studies develop methods for quality-adjustment of medical price indices (e.g. Cutler et al. 1998). As Hall (2017, p. 639) notes in her excellent review, "in general, the research literature shows adjusting for quality in the measurement of output in the medical sector to be quantitatively important." Yet most studies to date focus on acute episodes of care, leaving us with little evidence on the net value of spending on care for the leading cause of morbidity and mortality—chronic disease. Furthermore, few of these studies focus on Europe, and none to date focus on Asia, the most populous region in the world. We contribute to the literature by comparing the net value of care for an important chronic illness – diabetes mellitus – using detailed patient-level panel data on spending and health outcomes across four health systems: Japan, the Netherlands, Hong Kong, and Taiwan.

Our study applies a standardized methodology for estimating the net value of changes in medical spending, with population-specific risk prediction models to account for the over-prediction of cardiovascular events and death when western-based models are applied to Asian populations (Quan et al. 2019). This method measures value as the present discounted monetary value of any improved survival between the baseline and final periods, holding age and duration of diagnosis constant at their baseline values ("modifiable risk"), following Eggleston et al. (2009). In other words, modifiable risk is a decline in risk factors that is reasonably attributed to medical treatment. Our primary endpoint is modifiable risk of all-cause mortality for the next 5-year period; risk is based on biomarkers and is assessed with an appropriate risk model for individuals with type 2 diabetes for each race/ethnicity included. For survivors beyond the 5-year period, we assume remaining age- and sex-specific life expectancy based on the 2010 lifetable for each population.

From this measure of value, we subtract the increase in annual real modifiable spending per patient; this yields net value. Annual spending aggregates all health expenditures of the individuals diagnosed with diabetes, including out-of-pocket expenditures and spending on medications, without any attempt to disentangle diabetes-specific spending. This “cost-of-living” approach is analogous to a cost-effectiveness analysis comparing outcomes in the baseline and final periods, assigning a monetary value to life-years gained, and then subtracting the added costs of care (Eggleston et al. 2009).

The analyses draw from detailed individual-level panel data on medical utilization, spending, and clinical outcomes for patients in three health systems rooted in a Bismarckian tradition of social health insurance (Japan, the Netherlands, and Taiwan) and one Beveridge system (Hong Kong). More specifically, our sample includes 7,432 Japanese patients with diabetes, linked to biomarkers from mandatory annual health check-ups between 2010 and 2014; 14,312 patients with diabetes from a general practitioner (GP) registry in the Netherlands, linked to medical claims for 2008-2011; administrative and clinical data for 90,891 patients treated for diabetes with publicly-provided health services under the Hong Kong Hospital Authority between 2006 and 2014; and electronic medical records of 10,913 patients treated at a regional hospital in Taiwan from 2007 to 2013. Overall, with samples from four health systems spanning 2006 to 2014, we analyze individual-level panel data for 123,548 individuals with type 2 diabetes (5% of whom died during the study period) and more than 642,000 person-years of spending and outcomes.

Our results suggest both strong similarities across the four systems as well as the importance of quality-adjustment for understanding medical spending increases. Specifically, we find a positive net value of diabetes care, implying that the value of improved survival outweighs the added costs of care on average in each of the four health systems. For U.S.$100,000 value of a life-year, mean net value, standardized by age and sex, was $646 for Japan, $3,669 in the Netherlands, $3,985 in Hong Kong, and $10,717 for Taiwan. The net values remain positive even if one assumes only a value of $50,000 per life-year gained, which is analogous to assuming only half of survival gains valued at $100,000 were due to improved medical care.

This conclusion would only be strengthened, of course, to the degree the newer medical treatment improves non-fatal outcomes or quality of life. We therefore conclude that the increase in medical spending for management of diabetes is offset by an increase in quality, as measured by a reduction in risk of all-cause mortality.

Our finding that the quality-adjusted "cost of living" medical price index for managing diabetes has been declining across all four health systems is robust to various sensitivity analyses accounting for selective survival, end-of-life spending, and a range of values for a life-year or percentage of survival benefits attributable to medical spending. We discuss comparisons with earlier studies using this methodology in the U.S. (Eggleston et al. 2009) as well as related studies using different methods and studying other conditions (Cutler and McClellan 2001; Shapiro et al. 2001; Berndt et al. 2002; Highfell and Bernstein 2014; Dunn et al. 2018; Wamble et al. 2019). In sum, empirical evidence suggests that increased medical spending is "worth it" on average, although productivity surely could be further enhanced by increasing high-value care and reducing low-value care. Improving measures of net value will be critical to that effort.

In the next section, we overview the literature on quality adjustment of medical prices and the empirical precedents for this analysis, including the classic study of acute myocardial infarction (AMI) by Cutler et al. (1998). Subsequent sections describe our data, settings, and methods; present empirical results; and discuss our findings, robustness to multiple sensitivity analyses, and limitations. We conclude with a brief discussion of the implications for health economics research and health policy targeted at supporting high net-value healthcare innovation.

Quality-Adjusted Medical Prices and the Value of Medical Spending: Overview of the Literature

Our study is related to several strands of economics literature, from debate on methods for estimating the statistical value of a life, and the appropriate monetary value of a quality-adjusted life year (QALY) used in cost-effectiveness studies and health technology assessment, to macroeconomic models demonstrating the logic of higher medical spending in terms of the non-diminishing marginal utility of life extension (Hall and Jones 2007). However, increasing medical spending does not automatically deliver life extension or better quality of life, and the numerous market failures in the health sector (Arrow 1963) imply that market forces alone cannot guide social investments to their highest-valued use. Hence, estimating the net value of additional medical spending is of first-order importance for understanding productivity in the health sector and, as that sector grows, for understanding the value of overall economic activity, that is, deriving an accurate quality-adjusted price index for an entire economy. Moreover, as incomes rise, and demographic and epidemiologic transitions mature, leading causes of morbidity and mortality shift ever more towards chronic diseases such as cardiovascular disease and diabetes. Therefore quality-adjustment for changes in spending on chronic disease becomes the leading task for assessing productivity of health sectors.

Economic literature on the value of health improvement

The value of improvements in survival and quality of life matters for most economic outcomes – ranging from labor force participation (Bhattacharya and Lakdawalla 2006) to overall assessment of well-being (Frey and Stutzer 2002; Kahneman et al. 2004; Hall and Jones 2007; Deaton 2013). To illustrate, Jones and Klenow (2016) estimate that a consumption-equivalent measure of economic welfare that takes account of life expectancy and leisure shows a twenty-fold increase over the twentieth century rather than a seven-fold increase for per capita incomes alone. Murphy and Topel (2006) estimate the present value of longer life for a representative American at the turn of the twenty-first century, based on expected lifetime utility including the value of non-market time. They conclude that cumulative gains in life expectancy over the twentieth century were worth over $1.2 million to the representative American. If medical care accounts for even a modest fraction of that gain, medical spending increases would be “worth it” on average.

Accordingly, a central question in this literature is the extent to which health improvements can be attributed to medical care compared to other determinants -- social, genetic, environmental, and behavioral. Catillon et al. (2018) argue that personal medicine played a critical role in health improvements during the second half of the twentieth century but that medical care productivity may have decreased more recently. Studying the U.S. since 1960, Cutler et al. (2006) find that the increases in medical spending “provided reasonable value” but that spending increases on the elderly may have been less valuable than the average in recent decades.

The studies just cited utilize data primarily from the U.S.; it remains unclear to what extent their results generalize to other health systems in Europe or Asia. Does the value of medical innovation differ in single-payer systems or in those with a larger or smaller role of market forces in insurance and health service delivery? By comparing net value of a specific chronic disease across diverse health care and financing systems, our study contributes to answering this research question.

Value of new technologies, price changes, and international comparisons

Multiple studies document the health gains associated with specific technologies such as new pharmaceuticals (Lichtenberg 2019) or other new medical technologies (Hult et al. 2018). Others provide quantitative evidence on how specific medical treatments have reduced complications and major events from the leading chronic diseases, such as heart disease, stroke, and diabetes. Studying the causes of reduced medical spending growth among those age 65 and older in the U.S. during 1999–2012, Cutler and colleagues find that slower growth in spending for cardiovascular diseases accounted for about half of that slowdown, and that approximately half in turn of the reduction in major cardiovascular events was attributable to medications controlling cardiovascular risk factors (Cutler et al. 2019).

In a separate line of health economics research, careful empirical work documents medical price changes—differing across sectors, over time, and by condition—using actual negotiated prices paid by insurers in the U.S. (e.g. Newhouse and Garber 2013, Cooper et al. 2019) and internationally comparing differences in expenditures (e.g. Anderson et al. 2019). However, these studies tend to focus on the level of prices and spending across providers, regions, or OECD countries, without quality-adjusting price changes over time to examine the net value of expenditure increases.[1]

Related studies measuring productivity of healthcare providers do attempt to take account of quality change. For example, Romley et al. (2015) find that measured growth in hospital productivity switched from negative to positive when accounting for quality (as measured by health outcomes, adjusted for severity, among Medicare patients treated for heart attack, heart failure, and pneumonia during 2002–11).

[1] See comment by Baicker and Chandra (2018) on Papanicolas et al. (2018).

History of the cost-of-living approach

Before the 1990s, most research on medical spending and its "productivity" focused on measuring value with metrics of healthcare utilization or processes of care—thus focusing more on healthcare inputs or throughput, rather than on health outcomes. Research dating back to the pioneering work of Scitovsky (1965) explored a more theoretically-grounded approach that measured the value of health care spending not with inputs but with the desired output: cure or management of a specific medical condition (see Berndt et al. 2000 for a review). Although this approach is conceptually appealing, its empirical application has been bedeviled by multiple challenges of measurement and attribution. Data limitations often preclude measuring quality with health outcomes. Moreover, parsing both the outcomes and medical spending into components directly attributable to specific medical technologies and health conditions is challenging, even with detailed data. As a result, it is difficult to determine the extent to which health care improves health outcomes (Aizcorbe et al. 2018).

The earliest studies that attempted to account for health outcomes focused on acute conditions, where the value of a well-defined episode of medical spending can be measured by an unambiguous outcome—mortality. We adopt and modify this approach developed in the seminal work by Cutler et al. (1998). Their "cost-of-living index" measures the quality-adjusted cost of treating a health condition, using mortality to measure quality. Cutler et al. (1998) estimate that a quality-adjusted price index for AMI declined about 1 percent annually in the U.S. between 1983 and 1994. Similarly, in an update based on extending the same methods, Skinner et al. (2006) find that overall gains in post-AMI survival outweighed the 1986-2002 increase in AMI treatment costs.

Health economists applied this approach to several other medical conditions, corroborating the importance of quality adjustment for estimating the value of changes in medical expenditures. For example, Berndt et al. (2002) found a declining incremental cost of treating an episode of acute phase major depression between 1991 and 1996. Focusing on cataract surgery between 1969 and 1993, Shapiro et al. (2001) found that spending did not increase faster than the general price level. Cutler and McClellan (2001) presented evidence of positive net value for five conditions: AMI, low-birthweight infants, depression, cataracts, and breast cancer, with the first four showing a strongly positive net value, and breast cancer closer to zero (i.e. benefits and costs of similar magnitude). Highfell and Bernstein (2014) estimated changes in outcomes in the U.S. with disability-adjusted life years from the Global Burden of Disease in 1990 and 2010, linked to per-patient medical spending from the Medical Expenditure Panel Survey. They found positive net value overall and for the majority of diseases studied, with the largest health gains and net value for cardiovascular disease and cancer.

Previous studies on net value of diabetes care

Diabetes is an important chronic disease to study, given its growing prevalence, new treatment technologies, relatively complicated management, and life-long health implications. Recent U.S. evidence suggests that diabetes is one of the few medical conditions with significant increases in both numbers treated and per capita spending per treated individual (the latter increasing from $2309 in 1995 to $3402 in 2015; Biener et al. 2019).

Based on a 613-patient Mayo Clinic sample, Eggleston et al. (2009) found positive net value of spending on diabetes management in the U.S. between 1997 and 2005: $6,931-$10,911 for $200,000 per life-year, or $1,050- $2,215 for $100,000 per life-year gained. A non-trivial fraction of the overall net value (from $350 to $3,432, depending on the model) stemmed from avoided treatment spending for heart disease. In an extension of that study to the 1999-2009 period, Eggleston et al. (2011) similarly found a small positive net value, implying that quality improvements roughly offset the increased treatment costs on average, or that the quality-adjusted unit cost of care was not increasing. Wamble et al. (2019) estimate $6,377 cost per disability-adjusted life-year gained for U.S. diabetes patients, 1995-2015.

Our study contributes to this literature in several ways. First, when compared to previous studies, we assemble a larger dataset of patient-level panel data, linking medical spending to biomarkers. As noted above, Eggleston et al. (2009) was based on only 603 patients and Eggleston et al. (2011) on 821 patients. By contrast, we analyze individual-level panel data for 123,548 individuals with type 2 diabetes. Second, we explicitly compare results across different health systems of the high-income world, to see whether quality-adjusted price changes differ significantly across divergent health systems: a Bismarckian social insurance system in East Asia (Japan); a hybrid system of regulated competition with mandatory insurance, competing insurers, and strict regulations (Netherlands); a Beveridge-style government-run public system (Hong Kong); and a single-payer National Health Insurance system (Taiwan). Third, to adjust for quality changes appropriately, given different racial and ethnic differences in how biomarkers predict specific diabetes-related outcomes, we use risk prediction models tailored to the population in each sample. Finally, to explore robustness to various factors, we develop methods to account for decedents and undertake several sensitivity analyses. By building upon these methods, health economists can develop a richer basis for assessing the value of medical spending and comparing net value across even more diverse health systems and clinical settings.

Data and methods

This study draws on data from four health systems, three in Asia and one in Europe. We first provide an overview of the data and institutional setting for each, and then discuss our methods.

Data and institutional setting

Japan

Japan's social health insurance system provides universal coverage through insurance programs managed by employers for their employees, as well as insurance programs managed by municipalities (called National Health Insurance, NHI; see discussion in Sakamoto et al. 2018). The benefit package and payment systems are uniform nationally. A mixed-ownership service delivery system provides accessible care. Diabetes constituted the third largest disease category in 2014, with a national prevalence rate of 7.7% that is increasing as the population ages; more than one-quarter of Japan's adult population may have pre-diabetes or diabetes (Iizuka et al. 2017; Hashimoto and Eggleston 2019). According to the International Diabetes Federation data for 2017, Japan's total healthcare expenditure on its 7.2 million people with confirmed diabetes aged 20-79 was the fifth highest in the world (Iizuka et al. 2017).

We study the net value of care for patients with type 2 diabetes in Japan between 2010 and 2014, using medical claims and risk factors from mandatory annual health check-ups among 7,432 employed adults (aged 19-72) drawn from the database of employer health insurance claims provided by JMDC Inc (www.jmdc.co.jp/en). Individuals in our sample range from those newly diagnosed in 2009-2010 to those with duration of diagnosis over 30 years. The most reliably enforced mandatory health check-ups are for employees aged 40-74 working in large firms, such as the insured individuals in our dataset (Iizuka et al. 2017). Since our data comes from insurance plans managed by corporations, the insurer in this case is the employer. However, limitations on the proprietary data preclude linking any claims information to specific firms or measures of at-work productivity.

The Netherlands

Within the Dutch health care system, mandatory private insurance is financed by a mix of income-based, employer-paid, and privately-paid premiums. Insurers are obliged to accept all individuals for insurance and receive a case-mix-based payment to avoid cherry picking (van den Berg et al. 2011). Residents choose a health insurance policy with the insurer of their choice. The content of the basic insurance package is determined by the Minister of Health and consequently similar between insurers. Citizens may change their insurer on an annual basis, and about 6–8% of enrollees do so. In 2015, nearly 975 thousand persons aged 20-79 had diabetes in the Netherlands, which is a rate of 79 per 1000 (International Diabetes Federation 2017). Nearly 90% of this is type 2 diabetes, which is one of the three most prevalent chronic disorders. Ongoing aging and increases in the number of overweight people imply that the number of persons with diabetes will continue to rise. Primary care has been estimated to provide more than 80% of care for individuals with type 2 diabetes (Rodriguez-Sanchez et al. 2018).

We investigated the net value of diabetes care in a Dutch primary care setting. Clinical data from a Dutch GP-registry, Zodiac, were linked to an all-payer claims dataset from Vektis Health Care Information Center, covering the period 2008-2011. This linked dataset comprised nearly 16,000 individuals aged 40-90 with type 2 diabetes and included medical spending (limited to the mandatory insurance package and copayments) and diabetes-related outcomes including lab measurements. The difference between 2011 and 2008 complications risk was valued in monetary terms and predicted medical spending (in 2010€ and corrected for Elixhauser comorbidities, age and diabetes duration) was subtracted (Rodriguez-Sanchez et al. 2018). Risks were estimated using the UKPDS risk engine, as described below in the methods section.

Hong Kong

Hong Kong is the most developed among Chinese cities. Given its economic trends, lifestyle changes, and current stage of population aging, Hong Kong can arguably serve as a sentinel for the future of Mainland China. The majority (94%) of the population is ethnically Chinese, mostly second- or third-generation immigrants from the southern Chinese province of Guangdong. From 2006 to 2014, overall prevalence of diabetes standardized by age and sex in Hong Kong increased from 7.2% to 10.3% (Quan et al. 2017). Age- and sex- adjusted incidence slightly declined from 10.0 per 1,000 person-years in 2007 to 9.5 per 1,000 person-years in 2014. The overall prevalence of pre-diabetes was 8.9% in 2014. The prevalence of diabetes and pre-diabetes is expected to further increase as the population ages.

Hong Kong has a dual-track public and private health care system. Public-sector services, provided by the Hospital Authority, account for the majority of inpatient care (90% of total bed days and 80% of admissions), 50% of specialist outpatient care, and 30% of first-contact outpatient services. Universal coverage is provided by the public-run system with the government subsidizing over 90% of funding. In our net value analyses, we use the Hospital Authority electronic health record data on people with diabetes treated in the public health care system, which includes socio-demographic information, clinical and laboratory records, health service utilization, and prescribed medications.

Taiwan

Taiwan's population of 23 million has enjoyed universal health coverage since 1995 under the single-payer National Health Insurance (NHI) program, which integrated three pre-existing social insurance schemes and extended coverage to the remaining uninsured. Taiwan NHI offers comprehensive benefit coverage of ambulatory and inpatient services, delivered by a mix of public and private providers within Taiwan's market-oriented health care delivery system. The prevalence of diabetes in Taiwanese adults (aged 18 and above) was estimated to be 12.4% in 2013-2015 (Taiwan Health Promotion Administration, 2016). Diabetes treatment cost accounted for 4.24% of total NHI expenditures, next to renal disease and diseases of oral cavity and salivary glands in 2017 (Taiwan NHIA, 2018).

Our analytical dataset, which includes detailed clinical indicators, medical expenditures and sociodemographic information, includes individuals diagnosed with and treated for type 2 diabetes in the outpatient and/or inpatient departments of a large regional hospital in northern Taiwan from 2007 to 2013. These patients include participants and non-participants in a pay-for-performance program for diabetes management (Chen et al. 2016; Hsieh et al. 2015; Cheng et al.

2012; Chiu et al. 2019). A separate analysis focuses on comparing the net value of the P4P program with usual care (Lu et al. 2019); this analysis uses the entire sample. The selection criteria were that an individual was either enrolled in the diabetes P4P program or had no fewer than 3 diabetes visits in a year in the baseline period.

Methods

We apply and extend the methods developed by Eggleston et al. (2009) to estimate the net value of spending on diabetes care. Because the health benefits of medical treatments often manifest over long periods, those methods require prediction of future outcomes based on biomarkers from individuals in each time period, linked to utilization and spending. Then we can assess the monetary value of the change in quality—as proxied by predicted risks—between the baseline and final periods, net of change in spending. Importantly, estimating the net value requires risk prediction models appropriate to the characteristics of a given population.

Using patient biomarkers from our four samples, we estimate 5-year risks of all-cause mortality.[2] Following the methods of the previous Mayo clinic study, we define "modifiable" risks by holding at the baseline value the patient's age, and for the Japan, Netherlands, and Hong Kong samples, diabetes duration. For example, for a 60-year old woman who in 2010 had been living with diabetes for 4 years, we calculated baseline predicted risk using age 60, duration 4 years, and other risk factors such as systolic blood pressure as it was measured in 2010. We then calculated her modifiable risk in 2014 using her 2014 blood pressure, HbA1c, and other risk factors, but assuming her age in 2014 was still 60 and duration of diagnosis was still 4 years. Our net value calculations are based on these modifiable risk estimates to isolate the change in quality that is plausibly attributable to clinical care (including admonishments from providers to stop smoking, exercise more, etc.), rather than the natural process of aging and number of years the patient has lived with diabetes.

Our primary endpoint is modifiable risk of all-cause mortality for the next 5-year period based on the appropriate risk model for each race and ethnicity. For survivors beyond the 5-year

---

[2] One of our robustness checks also predicts the risk of major diabetes complications, using the Japanese sample and the Japan-specific risk prediction engine developed by Tanaka et al. (2013).

period, we assume remaining age- and sex-specific life expectancy based on the relevant 2010 lifetable.

We measure quality or value as the present discounted monetary value of any improved survival, holding age and duration of diagnosis constant at their baseline levels (the aforementioned "modifiable risk"). Net value subtracts the increase between the baseline and final periods in annual real modifiable spending per patient, the estimation of which is described below. Annual spending aggregates all health expenditure of the individuals diagnosed with diabetes, including out-of-pocket expenditures and spending on medications, without any attempt to disentangle diabetes-specific spending from non-diabetes spending.

Denote the 5-year risk of all-cause mortality for individual j in period t as P_{jt}. We calculate the value of reductions in fatal risk as follows: Assuming a 3% annual real discount rate applied halfway through each year and a given value of 1 life-year ranging from $50,000 to $150,000, we estimate the present discounted value of remaining life for 2 time points: the baseline and final observation periods. To do so, we approximate the predicted probability of death in the next 5 years by giving one fifth of the predicted probability (0.2 P_{jt}) to each of the first 5 years, and assume that all patients surviving beyond year 5 (with probability 1- P_{jt}) have the same age- and sex-specific remaining life expectancy as a general individual from the lifetable of that population.

For comparability across our four samples, we standardized by age and sex to the World Health Organization (WHO) World Standard Population (see Appendix). Throughout our analyses, 95% confidence intervals are bootstrapped using the percentile method with 1,000 repetitions. We used each economy's GDP deflator as a measure of the opportunity cost of the overall economy to adjust value and spending to constant 2010 PPP U.S. dollars.

We estimated predicted spending to smooth variation in individual observations of actual spending, as well as to calculate modifiable spending (holding age and duration of diabetes at baseline values). We estimated the total predicted medical spending for the first and last period of the study for each patient. To account for differences in how risk factors affect spending, we estimated generalized linear models and assumed a log link and gamma distribution, as suggested by Manning and Mullahy (2001). All medical and pharmaceutical spending are included in the dependent variable. The independent variables are age, sex, and indicator variables for presence of Elixhauser comorbid conditions at baseline. For the samples having information on duration of diabetes, duration (years with diagnosis) is included, as well as an interaction term of age with duration. For samples lacking the duration of diabetes variable, age-squared is included. By using the estimated coefficients from the model, we predict total spending for each patient in the baseline and final periods, and then predict total spending while holding age (and duration of diagnosis) constant at their baseline values. The increase in predicted spending between the baseline and final periods is then compared with the value of health status improvement to calculate the net value of the additional spending.

Quality adjustment with health outcomes: Risk prediction models

Measurement of health outcomes utilizes a risk prediction model suitable for estimating all-cause mortality in each of our samples. When available, population-specific risk models are preferable because they more accurately reflect the contribution of current clinical management of specific risk factors for that patient population. For example, using the UK Prospective Diabetes Study (UKPDS), as in the U.S.-based studies (Eggleston et al. 2009 and 2011) would not have been appropriate for all our samples, because the UKPDS over-predicts coronary heart disease (CHD) for Asian populations (Quan et al. 2019), for whom stroke risk is relatively high (GBD 2016 Lifetime Risk of Stroke Collaborators). Other risk prediction models such the Risk Equations for Complications Of type 2 Diabetes (RECODe) have also been calibrated for Western populations and have been demonstrated to be less accurate for East Asian populations (Quan et al. 2019). Below we provide a brief overview; for more details on the risk factors used in each model (Appendix Table 1) and on how each risk model was estimated, see the Appendix.

Measurement of health outcomes for the Japan sample, including all-cause mortality, is based on an extension of the 5-year risk of developing major complications and mortality as predicted by the Japan Diabetes Complications Study/the Japanese Elderly Diabetes Intervention Trial risk engine (JJRE) (Tanaka et al. 2013). The JJRE incorporates 11 risk factors to predict macro- and microvascular complications among Japanese patients with diabetes (without diabetes complications except mild retinopathy): sex, age, HbA_{1c}, years after diagnosis, body mass index (BMI), systolic blood pressure, non- high-density lipoprotein (HDL) cholesterol, albumin-to-creatinine ratio, atrial fibrillation, current smoker, and leisure-time physical activity. The model was developed based on data from 1,748 Japanese type 2 diabetic patients pooled from two clinical trials. Fortunately, Japan's mandatory annual health check-ups include a survey of various health behaviors including two questions that allow us to code leisure time physical activity.

For the Dutch sample, the UKPDS risk engine was applied. This model has been validated in the Dutch population and was shown to overpredict cardiovascular events (see discussion in Feenstra et al. 2019). The UKPDS risk engine was used to predict risk of fatal CHD and cerebrovascular events and then these risks were added to calculate total risk of fatal events as an approximation of diabetes-specific mortality risk. Age-specific all-cause mortality risks were then added, after correcting these for stroke and cardiovascular mortality, to avoid double counting. Risks of death from microvascular diabetes complications were considered negligible in comparison to these risks and ignored (Rodriguez-Sanchez et al. 2018). Variables required by the risk prediction model were taken from the linked dataset. Mortality was based on the VEKTIS data. Comorbidities were taken from coding available in ZODIAC.

Analyses for Hong Kong were based on risk prediction scores developed by Quan et al. (2019) for ethnic Chinese individuals residing in high-income East Asia by analyzing detailed clinical data for more than 1 million individuals over age 20 with type 2 diabetes in Hong Kong and Singapore. Based on more than 6 million person-years of follow-up between 2006 and 2016, the authors relate risk factors in a given year to the probability of mortality in the next five years. The Hong Kong University – Singapore (HKU-SG) model was developed using Cox proportional hazard models with the data for 678,750 participants from Hong Kong, and then validated with data for 386,425 participants from Singapore. The resulting HKU-SG risk score for mortality uses the following risk factors: age, duration of diabetes, gender, smoking, BMI, systolic and diastolic blood pressure, HbA1c, low-density lipoprotein (LDL) cholesterol and pre-existing conditions: atrial fibrillation, chronic kidney disease, ischemic heart disease and cerebrovascular disease (https://jquan.shinyapps.io/riskmodel/).

The Taiwan diabetes risk prediction model was constructed based on a sample of 18,202 type 2 diabetes patients treated in a large regional hospital in north-eastern Taiwan from 2007 to 2013. The national death registry was linked to the data to identify mortality and construct an all-cause mortality risk prediction engine using 8 risk factors in a Cox proportional hazard model: age, gender, history of cancer, history of hypertension, any use of anti-hyperlipidemic drugs, HbA1c, creatinine, and LDL/HDL ratio (Chiu et al. 2019). The Taiwan diabetes risk prediction model was internally validated based on the method in Chiu et al. (2017).

Sensitivity analyses and robustness checks

We undertake several sensitivity analyses. First, we compare net value for different assumptions about the value of a life-year, both using each sample's own age-sex structure and when standardizing by age and sex. We also compute results for different age-sex groups relative to age 60-64 for each sample, and net value stratified by duration quintiles (for the samples with diabetes duration).

We further modify and extend our methods with several robustness checks. The first examines the impact of including avoided treatment spending for major complications of diabetes. Using the JMDC Inc. claims data, the Japan team looked for people with a first stroke or development of CHD, and estimated associated expenditures in the incident year and incrementally for every subsequent year living with those complications. Then the present discounted value of avoided treatment spending was estimated based on reduction in the individual's modifiable predicted risks of CHD and stroke.

A second robustness check uses Cutler et al. (2006)'s assumption that medical care accounts for half of health gains, with non-medical factors playing an equally important role in health improvement. We compare our main results about net value assuming $100,000 per life-year gained with those using half that value (i.e., $50,000 per life-year). Our main results prove robust to assuming that only half of survival gains were due to medical care.

Because the main results about net value are based on those who survive to the final period, we undertake a third robustness check to account for decedents symmetrically with survivors. This check is especially important for our samples from Hong Kong and Taiwan, which are much larger and older than the working-age population analyzed by Eggleston et al. (2009 and 2011). As a result, medical care may be less effective at preventing mortality. The Netherlands sample excluded decedents.

To take account of selective survival and incorporate decedents and all their medical spending for the 3 samples with decedents, we estimate the net value of decedents, NV_D; this receives a weight in the overall net value that is proportional to the fraction of the baseline cohort that dies before the final period. NV_D is estimated in a different way from that of survivors, because it is based on cumulative spending and cumulative life-years, rather than subtracting baseline from final period values. To calculate NV_D we add the value of a life-year for every year survived, net of spending. Since even a small value of a life-year (e.g. $25,000) is much larger than mean expenditures in any given year in our samples (no more than $5000), this method yields a much larger net value than our main specification based on *change in* modifiable risk and expenditures between the baseline and final years. The net value for survivors receives a weight in the overall net value that is proportional to the fraction of the baseline cohort that survives to the final period in each sample. For more details, please see the appendix.

Results

Characteristics about the health systems and patient samples we analyze are provided in Table 1. All four represent high-income, mostly urban settings, with some of the highest life expectancies in the world, marked population aging, and growing burden of chronic diseases such as diabetes. Average age ranged from 52 in the Japanese working-age sample to 67.8 years in the Netherlands sample, with all but the Japan sample including individuals into their 90s. Most are balanced by gender, except that only one in five in the Japanese sample are women. Co-morbidity profiles also vary, with a more complicated clinical profile (e.g. 14% with renal failure) among patients regularly cared for at the regional hospital in Taiwan. This heterogeneity enriches our study of how quality-adjusted spending differs across different populations in different health systems, but also presents challenges for comparisons. For net value, we focus on age- and sex-standardized results.

Baseline average annual medical spending per individual with diabetes, in 2010 US$, varied across the four health systems: $2,526 in Japan (2010-11), $5,587 in the Netherlands (2008), $1,135 in Hong Kong (2006-08), and $2,474 in Taiwan (2007-09); see Table 2. The average increase in per-patient annual modifiable expenditures ranged from $108 more per year in the Netherlands between 2008 and 2011, to $896 more per year in Hong Kong between 2006-

08 and 2012-14. These trends (Figure 1) illustrate the relentless pressure of increasing input costs and technological change in medical care, even when removing the impact of aging and longer duration for those living with diabetes.

The increase in medical expenditures was accompanied by decline in modifiable risk of all-cause mortality (Table 2). The change in mean predicted modifiable mortality risk between baseline and final periods was negative (lower mortality) for all four samples, ranging in magnitude from one-fifth of a percentage point in the Netherlands and Taiwan to slightly more than a percentage point in Hong Kong. The percentage reduction in modifiable mortality risk was unsurprisingly largest in the sample with the longest panel of patients, Hong Kong, with a 10.67% reduction in age- and duration-constant risk of death. The percentage reductions in modifiable mortality risk were 8.96% in Japan, 2.70% in the Netherlands, and 1.32% in Taiwan.

Figure 2 shows the changes in mortality risk across age groups (in either age quintiles or 5- or 10-year groupings, depending on sample sizes) for each of the four systems. Across all age groups, modifiable mortality risk declined, with the largest absolute declines among the older individuals with larger absolute risk of death. These larger gains in modifiable risks among the oldest old may be driven not only by technological and behavioral change for better diabetes management, but also partially by selective survival. Our robustness check including decedents, described below, reveals that the overall pattern of improved outcomes is not driven entirely by selective survival. Indeed, accounting for decedents increases the magnitude of net value.

Because outcomes improved, quality-adjusted spending increases will be lower than unadjusted spending increases. The question remains, however, whether the net value of treatment spending was positive. Was the increase in spending associated with a commensurate or greater value of improved outcomes?

To answer this question, Table 3 reports the mean net value for each health system, for a range of life-year values. For comparability, the net values reported in Table 3 are standardized to the same age-sex composition of individuals, using the WHO world standard population. This standardization places a higher weight on younger cohorts and thus lowers the net value relative to almost all the samples' own age-sex composition of individuals with diabetes (see Appendix). Nonetheless, the average standardized net value for all four systems is positive. If life-years are valued at $100,000, mean net value ranges from $646 for Japan, to over $10,000 for Taiwan, with the Netherlands and Hong Kong in a similar middle range at $3,669 and $3,985, respectively. The net values are lower but remain positive when assuming life-years are valued at $50,000, or, equivalently, assuming that half of survival gains valued at $100,000 per life-year were attributable to medical care. Table 3 also shows the confidence intervals; every estimate is statistically significantly greater than zero.

Since the net value was positive, we conclude that the increase in medical spending for management of diabetes was offset by the value of improved quality, as measured by reduction in risk of all-cause mortality. Accordingly, the quality-adjusted "cost-of-living" medical price index for managing diabetes has been declining across all four health systems, to varying extents. These results are consistent with other studies estimating a modest positive net value for management of diabetes in the U.S. using similar methods (Eggleston et al. 2009, 2011) as well as contrasting methods (Wamble et al. 2019).

This finding of positive net value is robust to a range of robustness checks. The primary net value results (Table 3) derive from biomarkers in the baseline and final periods for each individual, and thus cannot be estimated for those who died before the final period. To account for decedents, we calculated an alternative measure of net value based on spending and outcomes per year survived. To illustrate, consider the Hong Kong sample of 90,891 individuals with diabetes, with 5% decedents; cumulative spending of all decedents (including spending in their year of death) averaged $2,488 per decedent-year, with a substantially higher amount spent in the last year of life (averaging $14,736 in 2011, or more than 7 times the average spending of survivors of $2,029 in the final period, 2012-14). This robustness check assigns the monetary value of a life-year to each year survived in the study period for all survivors and all decedents, net of their medical spending in all years survived. For this exercise, we use a conservative $25,000 value of a life-year. This is equivalent to attributing to medical care half of survival valued at $50,000 per life-year, or one-quarter of survival valued at $100,000 per life-year. We find that net value remains positive with this approach: the value of cumulative years survived outweighs cumulative spending among both survivors and decedents. Indeed, this method yields a much larger net value per individual with diabetes: $85,732 per diabetic for Japan, $141,969 per diabetic for Hong Kong, and $107,001 per diabetic for Taiwan. (The Netherlands sample did not include decedents). The appendix provides more details about the methods for this robustness check.

Taking account of avoided treatment spending

Improvements in medical treatment may also delay the onset of major complications such as stroke and heart disease, leading to improvements in quality of life and reduced spending on treatment for those conditions. In a robustness check with the Japanese data, we explored including the value of avoided treatment spending alongside improvements in survival.

One variant used an older sample and methods directly parallel to the original Mayo Clinic study. This analysis of JMDC Inc. data between 2008 and 2012 included a sample of 4,209 working-age adults with type 2 diabetes. Quality was measured in the same way as in the primary analyses: patient-level biomarkers in the baseline and final periods were inserted into the JJRE risk equations to predict 5-year risk. But in this case, we included not only risk of death but also risk of stroke and CHD. Any reduction in predicted modifiable risk of stroke or CHD was multiplied by the spending associated with those conditions (higher in the incident year and incremental for every subsequent year living with those complications) to estimate avoided treatment spending. We also employed the conservative assumption that net value would only be positive if the value of improved survival and avoided treatment spending more than offset actual treatment spending (not predicted modifiable spending).

This robustness check found that control of cardiovascular risk factors improved between 2008 and 2012, with a decline in 5-year modifiable risk of stroke (from 6.0% to 5.8%) and in modifiable fatal risk (from 2.2% to 2.1%) arising primarily in older age cohorts. Mean annual inflation-adjusted spending increased 12% overall and for all cohorts except those aged 60-69. For the overall sample of 4,209 individuals, the value of avoided stroke spending offset 74% of the increase in medical spending; adding in the value of improved survival (and a very small value from reduced CHD spending) offset the increase in actual spending entirely, yielding a slightly positive net value of $45 (at $100,000 per life-year; see Appendix Table 2). These results reinforce our conclusion that quality improvements have been at least as valuable as spending increases, even when using actual spending rather than predicted modifiable spending; in other words, a quality-adjusted price index was relatively constant across 2008-2012.

Analysis using the 2010-2014 larger Japanese sample also suggests that net value would be higher if we had incorporated non-fatal outcomes. Modifiable risk of stroke declined overall and for each quintile in the distribution of duration of diabetes, with declines of 14% to 17% depending on the duration. Similar reductions were observed in modifiable risk of CHD (11-16%) and non-cardiovascular mortality (3-10%). These improvements are sizable compared to the slight decrease in modifiable risk of all-cause mortality of -0.0018 percentage points from a baseline risk of 2.01% (i.e., 8.96% reduction in modifiable risk of death, see Table 2). Thus, while the age-sex-standardized net value in the Japanese sample was the smallest among our four health systems, incorporating the value of quality other than mortality reduction would substantially raise the net value, and appears to be particularly important when studying a working age population. A reduction in complications (or compression of morbidity through delay in development of complications to older ages) could also contribute to improved work productivity as well as avoided treatment spending.

Discussion

Analyzing over 600,000 patient-years of medical utilization, spending, and biomarker-based health outcomes, we find a small but statistically and economically meaningful positive net value of spending on diabetes management across four distinct health systems. This finding implies that the value of better outcomes more than offset spending increases, confirming that quality-adjustment matters quantitatively for understanding the value of medical spending.

Our estimates of net value are consistent with, and slightly higher than, those found for the U.S. (Eggleston et al. 2009 and 2011) despite a series of conservative assumptions. First, we use a lower range of values for a life-year ($50,000 - $150,000) compared to $200,000 used by the Mayo Clinic studies for the preferred specification. Our estimates are equivalent to attributing to medical care only a fraction of survival valued at $200,000 per life-year, with that fraction ranging from one-quarter to three-quarters. Second, we do not account for better quality of life or reduced probability of non-fatal adverse endpoints, such as blindness or amputation. Moreover, except for the sensitivity analysis for the Japan sample, our estimates do not include treatment spending avoided through management of those non-fatal complications. Third, we included all medical spending, with no attempt to isolate the fraction of spending directly attributable to diabetes. Fourth, our age-standardized sample places greater weight on younger ages than is typical of individuals with diagnosed diabetes, yielding a conservative estimate because net value is generally increasing with age, at least until 85+ (Figure 3). Fifth, the estimates do not include the value of any better work productivity arising from good diabetes management, such as less absenteeism and better on-the-job productivity.

One consideration, however, cuts the other way; in estimating survival gains based on predicted risks and remaining life expectancy, each team used a lifetable for their general population; none had access to lifetables estimated for individuals with diagnosed diabetes, but they likely are at greater risk of mortality even after the five-year horizon we used. In the future developing and using diabetes-specific lifetables would be a desired refinement of the net value estimates, especially for samples with a large proportion of individuals age 80 or older.

Our robustness checks suggest that incorporating the value of secondary prevention would increase the net value of diabetes care on average. The value of health spending could also be improved by investing in cost-effective primary prevention of diabetes (Diabetes Prevention Program Research Group 2002), hypertension (Moran et al. 2016), stroke (GBD 2016 Lifetime Risk of Stroke Collaborators), and other chronic diseases, including through improvement in leading preventable risk factors such as smoking, obesity and lack of physical activity.

Several other studies assessed changes in spending on specific conditions or groups of conditions, often finding that accounting for quality change reduced the spending increase or led to a decline in the quality-adjusted cost of treatment (e.g. Lakdawalla et al. 2015; Wamble et al. 2019). Net value is similar to the "quality-adjusted cost of care" proposed by Lakdawalla et al (2015), which is "cost growth net of growth in the value of health improvements, measured as survival gains multiplied by the value of survival." In fact, net value is precisely its opposite: net value is positive when quality-adjusted cost decreases, and negative when quality-adjusted cost increases. For example, Lakdawalla et al. (2015) estimate that for colorectal cancer, drug cost per patient increased by $34,493 between 1998 and 2005, but value of health improvements offset most of that spending increase, resulting in the quality-adjusted cost of care increasing by $1,377. This would be a net value of -$1,377 for colorectal cancer between 1998 and 2005.

The importance of quality adjustment for medical spending, and its associated focus on disease-specific spending and outcomes, has also prompted the attention of relevant government statistical agencies. For example in the U.S., recommendations from multiple researchers (e.g. National Research Council 2010) played a role in the U.S. Bureau of Economic Analysis decision to develop the disease-specific Medical Care Expenditure Index (Dunn et al. 2018). Hall (2017) provides a review of the research for the next step: quality-adjusting price indexes for specific medical conditions; she includes a useful taxonomy of *process-based adjustments* and *outcomes-based adjustments*. Our net value approach falls in the latter, outcomes-based category.

Further studies using these same methods can help to provide quality-adjustment for such disease-specific expenditure accounts, and to probe determinants of relative net value in different settings, such as accountable care organizations or pay-for-performance programs (Lu et al. 2019). Such studies can probe some of the areas of innovation for value in chronic disease management that we were unable to address in this study. Those limitations include data that are not necessarily representative of all patients with type 2 diabetes in each health system; different years and age ranges for the different systems; risk prediction models for East Asia that are not yet widely validated; and inability to consistently account for the value of avoiding non-fatal complications. Future research on these topics with other datasets and across diverse heath systems would be valuable, as would be further refinement of methods to account for quality of life among both decedents and survivors.

Conclusion

We find a positive net value of spending on diabetes care across four health systems, slightly greater in magnitude than earlier estimates of net value for diabetes care in the U.S. In other words, the value of quality improvements outweighed additional medical spending on average. If further studies reveal similarly positive net value for spending on other prominent chronic diseases, then it would indicate that quality-adjusted medical prices for chronic disease have been declining. Even a constant quality-adjusted unit cost would suggest that productivity of healthcare for chronic disease has been improving, since input prices have been increasing. However, positive average net value does not imply efficiency of all marginal spending increases. Further efforts to measure and reward net value could contribute to innovations that reduce wasteful or low-value spending, support high-value innovations, and provide a better measurement of overall economic activity by understanding quality-adjusted productivity in the health sector.

References

Aizcorbe, A., Baker, C., Berndt, E. R., & Cutler, D. M. (Eds.). (2018). *Measuring and Modeling Health Care Costs* (Vol. 76). University of Chicago Press.

Anderson, G. F., Hussey, P., & Petrosyan, V. (2019). It's Still The Prices, Stupid: Why The U.S. Spends So Much On Health Care, And A Tribute To Uwe Reinhardt. *Health Affairs*, *38*(1), 87-95.

Arrow, K.J. (1963). Uncertainty and the Welfare Economics of Medical Care. *American Economic Review,* 53(5), 941–73.

Baicker, K., & Chandra, A. (2018). Challenges in Understanding Differences in Health Care Spending Between the United States and Other High-Income Countries. *JAMA,319*(10), 986-7. doi:10.1001/jama.2018.1152

Bhattacharya, J., & Lakdawalla, D. N. (2006). The labor market value of health improvements. In *Forum for Health Economics & Policy* (Vol. 9, No. 2). De Gruyter.

Berndt, E. R., Bir, A., Busch, S. H., Frank, R. G., & Normand, S. L. T. (2002). The medical treatment of depression, 1991–1996: productive inefficiency, expected outcome variations, and price indexes. *Journal of Health Economics*, *21*(3), 373-396.

Berndt, E. R., Cutler, D. M., Frank, R. G., Griliches, Z., Newhouse, J. P., & Triplett, J. E. (2000). Medical care prices and output. In *Handbook of health economics* (Vol. 1, pp. 119-180). Elsevier.

Biener, A. I., Decker, S. L., & Rohde, F. (2019). Source of Increased Health Care Spending in the United States. *JAMA,321*(12), 1147. doi:10.1001/jama.2019.0679

Catillon, M., Cutler, D., & Getzen, T. (2018). *Two Hundred Years of Health and Medical Care: The Importance of Medical Care for Life Expectancy Gains* (No. w25330). National Bureau of Economic Research.

Chen, Y., Lee, C. T., Lin, B. J., Chang, Y., & Shi, H. (2016). Impact of pay-for-performance on mortality in diabetes patients in Taiwan. *Medicine,95*(27). doi:10.1097/md.0000000000004197

Chen, B., K. Eggleston, & T. Iizuka. (2019). Adherence, Vertical Integration, and Behavioral Hazard. Working paper. Stanford APARC Asia Health Policy Program.

Cheng, S., Lee, T., & Chen, C. (2012). A Longitudinal Examination of a Pay-for-Performance Program for Diabetes Care. *Medical Care,50*(2), 109-116. doi:10.1097/mlr.0b013e31822d5d36

Chiu SY et al. (2017). "Faecal haemoglobin concentration influences risk prediction of interval cancers resulting from inadequate colonoscopy quality: analysis of the Taiwanese Nationwide Colorectal Cancer Screening Program", Gut. 2017 Feb;66(2):293-300. doi: 10.1136/gutjnl-2015-310256.

Chiu, Sherry Yueh-Hsia, Ying Isabel Chen, Jui-fen Rachel Lu, Sheng-Fong Kuo, Chih-Hung Chen. (2019). Developing Predictive Model on All-cause Mortality Risk for Type 2 Diabetes Using Hospital-based Prospective Cohort Study. Working paper.

Cooper, Z., Craig, S., Gaynor, M., Harish, N. J., Krumholz, H. M., & Van Reenen, J. (2019). Hospital Prices Grew Substantially Faster Than Physician Prices For Hospital-Based Care In 2007–14. *Health Affairs*, *38*(2), 184-189.

Cutler, D. M., Ghosh, K., Messer, K. L., Raghunathan, T. E., Stewart, S. T., & Rosen, A. B. (2019). Explaining The Slowdown In Medical Spending Growth Among The Elderly, 1999–2012. *Health Affairs*, *38*(2), 222-229.

Cutler, D. M., McClellan, M., Newhouse, J. P., & Remler, D. (1998). Are medical prices declining? Evidence from heart attack treatments. *The Quarterly Journal of Economics*, *113*(4), 991-1024.

Cutler, D. M., & McClellan, M. (2001). Is technological change in medicine worth it? *Health affairs*, *20*(5), 11-29.

Cutler, D. M., Rosen, A. B., & Vijan, S. (2006). The value of medical spending in the United States, 1960–2000. *New England Journal of Medicine*, *355*(9), 920-927.

Deaton, A. (2013). *The great escape: health, wealth, and the origins of inequality*. Princeton University Press.

Diabetes Prevention Program Research Group. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England journal of medicine*, *346*(6), 393-403.

Dunn, A., Grosse, S. D., & Zuvekas, S. H. (2018). Adjusting health expenditures for inflation: a review of measures for health services research in the United States. *Health services research*, *53*(1), 175-196.

Eggleston, K. N., Shah, N. D., Smith, S. A., Wagie, A. E., Williams, A. R., Grossman, J. H., ... & Newhouse, J. P. (2009). The net value of health care for patients with type 2 diabetes, 1997 to 2005. *Annals of internal medicine*, *151*(6), 386-393.

Eggleston, K. N., Shah, N. D., Smith, S. A., Berndt, E. R., & Newhouse, J. P. (2011). Quality adjustment for health care spending on chronic disease: Evidence from diabetes treatment, 1999-2009. *American Economic Review papers and proceedings*, *101*(3), 206-11.

Frey, B. S., & Stutzer, A. (2002). What can economists learn from happiness research? *Journal of Economic literature*, *40*(2), 402-435.

GBD 2016 Lifetime Risk of Stroke Collaborators. (2018). Global, Regional, and Country-Specific Lifetime Risks of Stroke, 1990 and 2016. *New England Journal of Medicine*, *379*(25), 2429-2437.

Hall, Anne E. (2017). Adjusting the Measurement of the Output of the Medical Sector for Quality: A Review of the Literature. *Medical Care Research and Review* 74.6: 639-667.

Hall, R. E., & Jones, C. I. (2007). The value of life and the rise in health spending. *The Quarterly Journal of Economics*, *122*(1), 39-72.

Hashimoto, C., & K. Eggleston. (2019). Healthy Aging and Economic Research on the Net Value of Non-communicable Disease Management in Japan. Chapter 3 in *Healthy Aging in Asia* (K. Eggleston, editor), Stanford FSI Shorenstein Asia-Pacific Research Center publication series, Brookings Institution Press, forthcoming.

Highfell, T., & E. Bernstein. (2014). Using Disability Adjusted Life Years to Value the Treatment of Thirty Chronic Conditions in the U.S. from 1987-2010. Available at http://www.bea.gov/system/files/papers/WP2014-9.pdf

Hsieh, H., Gu, S., Shin, S., Kao, H., Lin, Y., & Chiu, H. (2015). Cost-Effectiveness of a Diabetes Pay-For-Performance Program in Diabetes Patients with Multiple Chronic Conditions. *Plos One,10*(7). doi:10.1371/journal.pone.0133163

Hult, K. J., Jaffe, S., & Philipson, T. J. (2018). How Does Technological Change Affect Quality-Adjusted Prices in Health Care? Systematic Evidence from Thousands of Innovations. *American Journal of Health Economics*, *4*(4), 433-453.

Iizuka, T., Nishiyama, K., Chen, B., & Eggleston, K. (2017). Is Preventive Care Worth the Cost? Evidence from Mandatory Checkups in Japan. NBER Working Paper No. 23413.

International Diabetes Federation. (2017) IDF Diabetes Atlas 2017. International Diabetes Federation (IDF): Brussels, Belgium. Available at https://diabetesatlas.org/resources/2017-atlas.html.

Jones, C. I., & Klenow, P. J. (2016). Beyond GDP? Welfare across countries and time. *American Economic Review*, *106*(9), 2426-57.

Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. (2004). Toward national well-being accounts. *American Economic Review*, *94*(2), 429-434.

Lakdawalla, D., Shafrin, J., Lucarelli, C., Nicholson, S., Khan, Z. M., & Philipson, T. J. (2015). Quality-adjusted cost of care: a meaningful way to measure growth in innovation cost versus the value of health gains. *Health Affairs*, *34*(4), 555-561.

Lichtenberg, F. R. (2019). *How Many Life-Years Have New Drugs Saved? A 3-Way Fixed-Effects Analysis of 66 Diseases in 27 Countries, 2000-2013* (No. w25483). National Bureau of Economic Research.

Lu, Jui-fen Rachel, Ying Isabel Chen, Karen Eggleston, Chih-Hung Chen & Brian Chen (2019). Assessing a Taiwan pay-for-performance program for the care of diabetes patients from the net value perspective, working paper, April 2019.

Manning, W. G., & Mullahy, J. (2001). Estimating log models: to transform or not to transform? *Journal of health economics*, *20*(4), 461-494.

Murphy, K. M., & Topel, R. H. (2006). The value of health and longevity. *Journal of political Economy*, *114*(5), 871-904.

National Research Council (US) Panel to Advance a Research Program on the Design of National Health Accounts. (2010). *Accounting for Health and Health Care: Approaches to Measuring the Sources and Costs of Their Improvement*. Washington (DC): National Academies Press (US). Available from: https://www.ncbi.nlm.nih.gov/books/NBK53346/ doi: 10.17226/12938

Newhouse, J. P., & Garber, A. M. (2013). Geographic Variation in Health Care Spending in the United States. *Jama,310*(12), 1227. doi:10.1001/jama.2013.278139

Papanicolas, I., Woskie, L. R., & Jha, A. K. (2018). Health Care Spending in the United States and Other High-Income Countries. *Jama,319*(10), 1024. doi:10.1001/jama.2018.1150

Quan, J., Li, T. K., Pang, H., Choi, C. H., Siu, S. C., Tang, S. Y., Wat, N.M, Woo, J., Johnston, J.M. & Leung, G. M. (2017). Diabetes incidence and prevalence in Hong Kong, China during 2006–2014. *Diabetic Medicine*, *34*(7), 902-908.

Quan, Jianchao, Deanette Pang, Tom K. Li, Cheung Hei Choi, Shing Chung Siu, Simon Y. Tang, Nelson M. Wat, Jean Woo, Zheng Yi Lau, Kelvin B. Tan, & Gabriel M. Leung. (2019). Risk prediction scores for mortality, cerebrovascular, and heart disease amongst Chinese people with type 2 diabetes.

Rodriguez-Sanchez, Beatriz, Talitha Feestra, Henk Bilo, and Jeroen Struijs. 2018. The net value of Dutch primary care management of patients with type 2 diabetes, focusing on the elderly. Presentation at EuHEA July 2018.

Romley, J. A., Goldman, D. P., & Sood, N. (2015). U.S. hospitals experienced substantial productivity growth during 2002–11. *Health Affairs*, *34*(3), 511-518.

Sakamoto, H., Rahman, M., Nomura, S., et al. (2018). Japan Health Systems Review. *Asia Pacific Observatory on Health Systems and Policies, vol 8, No 1.* http://www.searo.who.int/entity/asia_pacific_observatory/publications/hits/hit_japan/en/

Scitovsky, Anne A. (1967). Changes in the Costs of Treatment of Selected Illnesses, 1951–65. *American Economic Review,* 57(5): 1182–95.

Shapiro, Irving, Matthew D. Shapiro, & David W. Wilcox. (2001). Measuring the Value of Cataract Surgery. In *Medical Care Output and Productivity*. Vol. 62, Studies in Income and Wealth, ed. David M. Cutler and Ernst R. Berndt, 411–38. Chicago: University of Chicago Press.

Skinner, J. S., Staiger, D. O., & Fisher, E. S. (2006). Is technological change in medicine always worth it? The case of acute myocardial infarction. *Health affairs*, *25*(2), w34-w47.

Taiwan Health Promotion Administration, 2016. https://www.hpa.gov.tw/Pages/Detail.aspx?nodeid=1090&pid=6426, accessed on April 14, 2019.

Taiwan National Health Insurance Administration, 2018. https://www.nhi.gov.tw/Content_List.aspx?n=806314145D8E1187&topn=CDA985A80C0DE710, accessed on April 14, 2019.

Tanaka, Shiro, Sachiko Tanaka, Satoshi Iimuro, etal., for the Japan Diabetes Complications Study Group and the Japanese Elderly Diabetes Intervention Trial Group. (2013). Predicting Macro- and Microvascular Complications in Type 2 Diabetes, *Diabetes Care*. 36(5): 1193-1199.

UK Prospective Diabetes Study (UKPDS) Group. (1998). Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *The Lancet* 352.9131: 837-853.

Van den Berg, M., Heijink, R., Zwakhals, L., Verkleij, H., & Westert, G. (2011). Health care performance in the Netherlands: Easy access, varying quality, rising costs. *Eurohealth*, *16*(4), 27.

Wamble, D., Ciarametaro, M., Houghton, K., Ajmera, M., & Dubois, R. W. (2019). What's Been The Bang For The Buck? Cost-Effectiveness Of Health Care Spending Across Selected Conditions In The US. *Health Affairs*, *38*(1), 68-75.

Table 1. Characteristics of the Four Health Systems and the Four Samples of Adults with Type 2 Diabetes Mellitus

	Japan	Netherlands	Hong Kong	Taiwan
Population, 2017, in millions	126.8	17.1	7.4	23.6
GDP per capita, 2016 USD PPP	$41,476	$51,320	$58,561	$50,400
Health expenditures per capita (2014, USI	$3,703	$5,694	$2,222	$1,435
Data	JMDC Inc. medical insurance claims data for working-aged adults employed at large corporations, linked to annual health check-up data.	Linked data from ZODIAC (routine data from primary care physicians) and VEKTIS (national claims data covering mandatory health care insurance).	Administrative dataset of all publicly provided health care services. Participants with complete data at baseline and final periods.	Electronic medical records from one large regional hospital in northern Taiwan.
Observations (N individuals)	7,432	14,312	90,891	10,913
Survivors (N)	7,077	14,312	86,344	9,739
Decedents (N)	355	were excluded	4,547	1,174
Baseline period (year(s))	2010-11	2008	2006-2008	2007-2009
Final period (year(s))	2013-14	2011	2012-2014	2011-2013
Baseline characteristics				
Age: mean, (min-max)	52.02 (20-74)	67.76 (40 - 92)	60.78 (15 -101)	62.58 (19 - 99)
Female (share of sample)	0.205	0.51	0.526	0.507
BMI	25.49	27.18	25.6	NA
Systolic blood pressure	129.34	138.54	135.79	73.77% using antihypertensive
Diastolic blood pressure	73.9	NA	74.92	NA
HbA1c	6.73	6.68	7.58	7.38
Duration of DM, years	3.66	5.64	5.82	NA
Selected Elixhauser co-morbidities (share of sample with each)				
Hypertension	0.38	0.14	0.36	0.54
Renal Failure	0.01	0	0.04	0.14
Peripheral vascular disorders	0.06	0.13	0.04	0.02
Congestive heart failure	0.05	0.06	0.05	0.08
Cardiac arrhythmias	0.06	NA	0.10	0.05
Valvular disease	0.02	NA	0.01	0.04
Chronic pulmonary disease	0.11	NA	0.04	0.09
Hypothyroidism	0.01	0.01	0.01	0.02

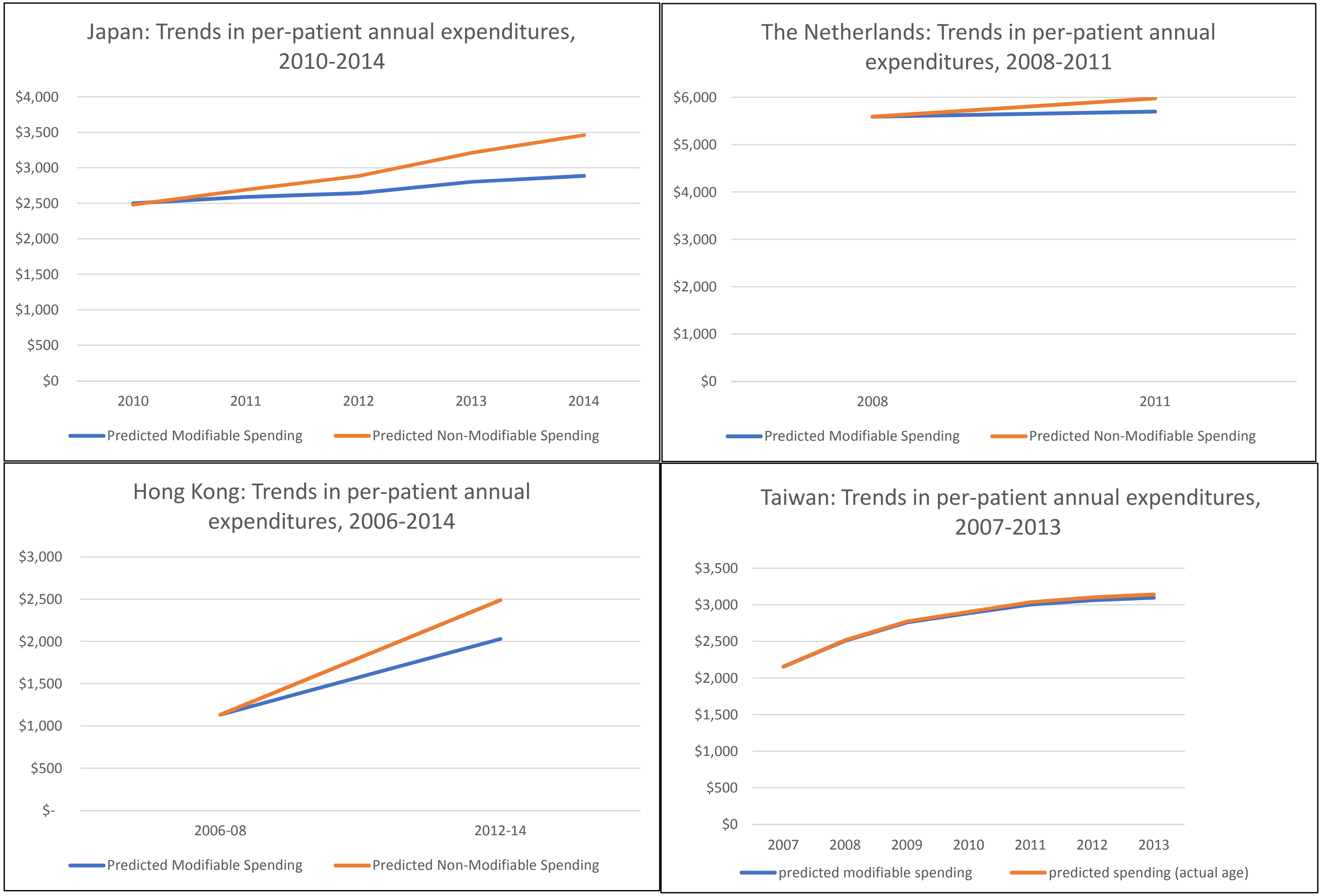
Sources: Authors' analyses of the four study samples; World Bank World Development Indicators. NA = Not available.

Table 2. Average change in spending and quality, as proxied by predicted mortality, holding age and duration constant

	Japan	Netherlands	Hong Kong	Taiwan
Predicted expenditures (baseline) (a)	$2,526	$5,587	$1,135	$2,474
Predicted modifiable expenditures (final) (b)	$2,813	$5,695	$2,029	$3,055
Change in annual expenditures (b-a)	$286	$108	$894	$580
Percentage change, baseline to final	11.34%	1.93%	78.71%	23.46%
Predicted mortality (baseline) (c)	2.01%	7.40%	9.93%	17.47%
Predicted modifiable mortality (final) (d)	1.83%	7.20%	8.87%	17.25%
Change in modifiable mortality risk (d-c)	-0.18%	-0.20%	-1.06%	-0.23%
Percentage reduction in modifiable mortality risk, baseline to final	8.96%	2.70%	10.67%	1.32%

Note: "Modifiable" refers to predictions of spending or mortality risk when the individual's age and duration of diagnosis are held constant at their values in the baseline period. Only other risk factors, such as control of blood sugar and blood pressure, differ. See text.

Figure 1. Change in per-patient annual real expenditures between baseline and final periods (2010 US dollars)



Note: Predicted modifiable spending is the spending predicted when holding age and duration of diagnosis constant at baseline values, while predicted spending used patients' actual age and duration of diagnosis in each year. Spending was converted to 2010 USD.

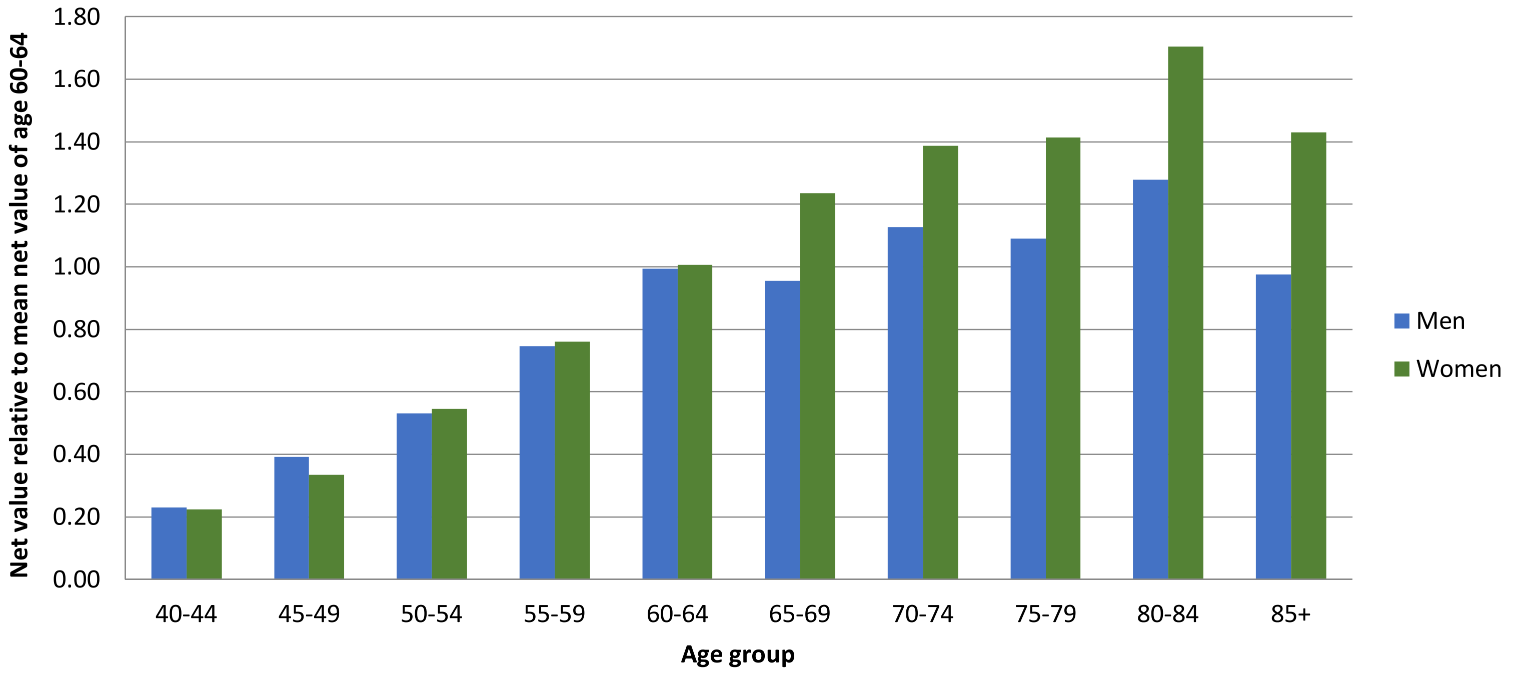
Figure 2. Quality trends: Change in modifiable mortality risk, by age category

Table 3. Mean net value of spending on diabetes management in four health systems, for a range of values of a life-year

Value of a life-year	Japan	Netherlands	Hong Kong	Taiwan
$50,000	**$259**	**$1,799**	**$1,874**	**$5,119**
Confidence interval	$220 - $297	$1,546 - $2,053	$1,785 - $1,965	$4,478 - $5,761
$100,000	**$646**	**$3,669**	**$3,985**	**$10,717**
Confidence interval	$570 - $722	$3,162 - $4175	$3,846 - $4,122	$9,437 - $11,996
$150,000	**$1,018**	**$7,407**	**$6,096**	**$16,254**
Confidence interval	$901 - $1,135	$6,649 - $8,167	$5,909 - $6,268	$14,357 - $18,151

Note: Standardized by age and sex according to the WHO world standard population.

Figure 3. Net Value by Sex and Age Group, Relative to Mean Net Value of Age 60-64 Individuals in Each Health System



Note: Each column represents the net value of that specifc age-sex-group relative to the mean (male+female) net value for age 60-64 adults in each health system, averaged with weights corresponding to sample size. Net values for age 70 and older exclude Japan, and thus represent weighted averages of the Netherlands, Hong Kong, and Taiwan samples, weighted by the number of patients in each sample.