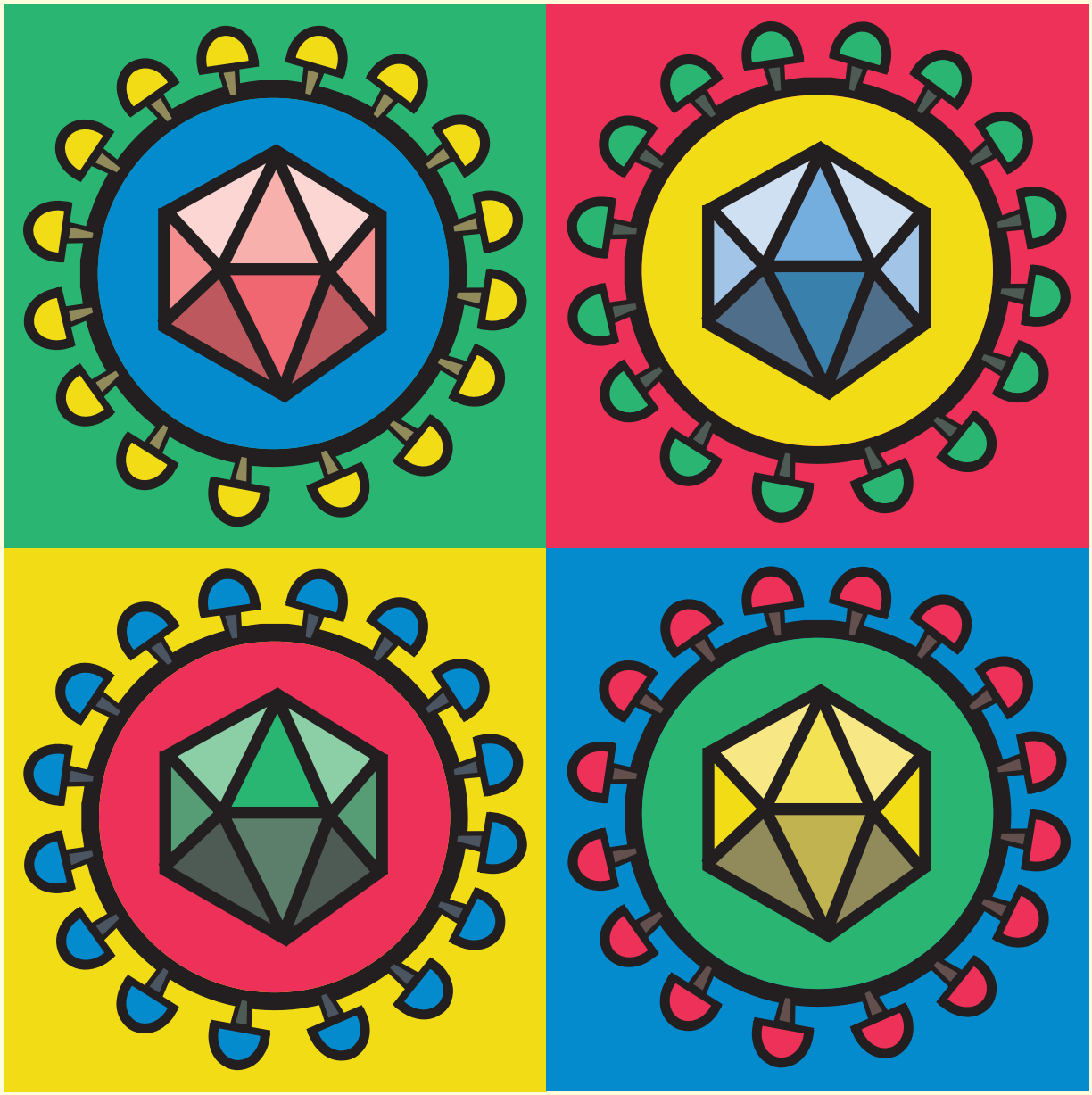
HMSR

Harvard Medical
Student Review

Issue 3
September 2015



HMSR

September 25th, 2015

To our esteemed readers,

It is my honor and pleasure to present the third issue of the Harvard Medical Student Review, a student-run, peer-reviewed journal that explores a rich diversity of topics related to medicine and healthcare.

As you may know, HMSR was recently inherited by a new leadership team. We hope to build upon the legacy of excellence and scientific inquiry established by our predecessors, as well as propel HMSR to new heights. As part of this campaign, we aim to strengthen HMSR's connections to academic and research institutions around the world, creating a global network for student-engaged scientific investigation.

Behind every successful venture is an exceptional team. We would like to thank our Associate Editors, Artists, Authors, and the previous HMSR Board for serving as an invaluable source of inspiration. In particular, we thank Arthur, Eric, Leigh-Ann, and Mark for their mentorship and guidance. Additionally, we express our deep appreciation for the unwavering support of our sponsors and partners. Finally, I would like to personally thank the 2015-2016 HMSR Executive Board members for their great enthusiasm, steadfast dedication, and tireless efforts.

Enjoy the issue. We look forward to hearing from you.

All the best,

Soraya Azzawi
Editor-in-Chief

About HMSR

The Harvard Medical Student Review (HMSR) is student-founded, student-managed, and student-administered under the guidance of faculty and staff. Its mission is to provide a platform for students to contribute to important issues facing health and medicine through a variety of formats, including scholarly articles, editorials, and original artwork. Contributions are invited from the Harvard medical, dental, and public health schools, the rest of Harvard University, and other medical schools.

The cover art was created by Kathy Wang to represent this issue's article, "Historical Path of Discovery of Viral Hepatitis" by Wong et al.

HMSR

Table of Contents

September 2015; Issue 3

ARTICLE

A Business Case for Oral Physicians: Market Analysis and Potential Practice Models for Dentists to Address the United States Primary Care Shortage 2–13
Sina Hedayatnia, BA, and Donald B. Giddon, DMD, PhD
Artwork: Jessie Scheepers *3*

REVIEW

Preventing Stroke in Atrial Fibrillation: Warfarin vs. Factor Xa Inhibitors 14–17
Pratyaksh K. Srivastava, BA
Artwork: Faizan Khan *15*

Historical Path of Discovery of Viral Hepatitis 18–36
David T. Wong; Martin C. Mihm, MD; James L. Boyer, MD; Dhanpat Jain, MD
Artwork: R. Frederick Lambert *23*

VIEWPOINT

Alma Mater Revisited: Teaching Medicine as a Fulbright Scholar 37–41
Gyorgy Baffy, MD, PhD
Artwork: "Tempus Fugit" by Erin Dvorak *39*

Increasing Drug Costs in the United States: Time for Reform? 41–45
Janani Panchalingam MPH; Madhu Mazumdar, PhD; Jashvant Poeran, MD, PhD
Artwork: Ryan Xiao *43*

The "Harm Principle" in the Context of Organ Donation 46–55
John Fortunato, MA
Artwork: "A Doctor's Dilemma" by Anas Elbashir Ahmed Musa *50*

HMSR

ARTICLE
Health Policy

A Business Case for Oral Physicians: Market Analysis and Potential Practice Models for Dentists to Address the United States Primary Care Shortage

Sina Hedayatnia, BA[1,2] and Donald B. Giddon, DMD, PhD[3]

[1]Harvard School of Dental Medicine, Boston, MA 02115, USA
[2]Harvard Business School, Boston, MA 02163, USA
[3]Department of Developmental Biology, Harvard School of Dental Medicine, Boston, MA, 02115, USA
Correspondence should be addressed to S.H. (Sina_Hedayatnia@hsdm.harvard.edu)

Dentists are in a prime position to address the US primary care shortage by expanding their scope of practice towards screening for systemic health conditions. Primary care activity by dentists alone has the potential to save $15 billion in avoided medical expenses through concierge models and/or ACO-partnered dental practices.

Abstract

The United States is experiencing a dramatic shortage of primary care providers expected to worsen over the next decade, primarily from population growth and health care reform. The current environment fails to adequately diagnose the nation's growing systemic health issues. Of the 80 million people affected by cardiovascular disease and 23 million people with diabetes mellitus, estimates suggest that one-third are unaware of their condition. Early screening for these and other conditions can elucidate risk factors for patients and help prevent disease through medical intervention and counseling.

One possible solution to the primary care gap is through the dentist's office. Dentists are trained in all manners of systemic illnesses, especially those that manifest in the oral cavity. Both medical and dental communities increasingly accept the concept of the oral physician, a dentist who can provide medical screening and limited primary care. In order for dental professionals to widely endorse their necessary role as oral physicians, it is important to develop a business model that not only addresses oral and systemic health, but also remains sustainable in the current health care marketplace.

Here, we conduct a market analysis for the oral physician by estimating the population with undiagnosed health conditions who receive regular oral health care but not medical care, estimating the total costs of non-intervention in the form of medical costs, and developing practical models for dentists to address primary care in their practices.

Introduction

The demand for primary care services in the United States is rising rapidly. Between 2005 and 2025, the number of patient visits to all primary care practitioners is expected to increase 29% due to population growth and an in increase in the number of elderly patients [1]. Under the Affordable Care Act, the population of insured patients seeking care is

Artist: Jessie Scheepers
Cape Town, South Africa
jessie.scheepers@meltwater.com |
www.meltwater.com

expected to increase the patient pool by an additional 5–8%. Meanwhile, the number of primary care providers is expected to grow only by 2–7%, with more medical students forgoing primary care tracks in favor of higher paying, less demanding specialty positions [2]. Compounding the problems associated with a paucity of providers is the misallocation of physicians, nurse practitioners, and physician assistants. Rural regions are generally the most underserved; these areas are served by 10% of the physician workforce, despite comprising 21% of the US population [2].

Coupled with the primary care shortage are millions of Americans who suffer from chronic systemic health conditions, often resulting in patients who fail to be properly diagnosed. 80 million Americans are recognized as having cardiovascular disease (CVD), while 23 million are affected by diabetes mellitus [3]. Despite their widespread prevalence and associated health risks, it is estimated that one-third of these diseases remain undiagnosed. These conditions can be prevented by early intervention provided that a medical screening takes place before the disease onset. Interventions following screening for such diseases, leading to dietary changes and physical activity, have led to significant decreases in disease incidence [3].

Numerous authorities in the field of dentistry have endorsed the concept of the "oral physician." An oral physician is a dentist or dental specialist with a complete knowledge of systemic health who can provide medical screenings and limited primary care. Dr. Donald Giddon, former dean of the New York University College of Dentistry, and Dr. Bruce Donoff, current dean of the Harvard School of Dental Medicine, among others, have supported dental curricula that go beyond the tooth and its supporting structures and go towards the pathophysiology and management of systemic diseases. Dentists and dental specialists are in a prime position to provide these services.

The connection between oral and systemic health is a topic of growing scholarly interest. Research in periodontal medicine is investigating effects of gum disease on cardiovascular and respiratory diseases, diabetes mellitus, and pregnancy [4]. Systemic diseases can also have oral manifestations, as in the case of hematological disease, cancer, and autoimmune diseases [4]. Interventions for smoking and diet management are already taking place in dental practices, given the role of cigarettes as a factor in oral squamous carcinoma and poor food choices as a risk factor for tooth decay. Some dentists consistently recognize and provide referrals for issues ranging from hypertension and skin cancer to domestic and substance abuse [5]. Given the two-way relationships between gum health and general health, periodontists, who are specialists in supporting structures of teeth, are well suited to address systemic health concerns in their care for patients. Orthodontists, who specialize in treating dental malocclusions, build strong relationships with teenaged patients, often seeing them once a month for 2 years or more. As such, they may be capable of screening and counseling adolescent patients on developmental, eating, and behavioral disorders [6].

An interesting avenue for the dental profession to address primary care is through mental health. In addition to oral manifestation of psychological ailments, such as ulcers arising from lowered oral immune defenses under neurohormonal control, dental and craniofacial abnormalities can lead to depression and anxiety, especially in cases of body dysmorphic disorder and abuse [7]. Interventions by a dental professional could be crucial in instances when a patient has been unaware of a problem or unwilling to seek out professional psychiatric care due to the stigma of mental health issues.

The dental community largely supports their rebranding as oral physicians. In a recent survey of US dentists, a majority of oral health professionals supported the incorporation of medical health screenings in the dental setting. 83.4% of dentists surveyed were willing to perform chair-side medical screenings, which would yield immediate results for diagnosis and management of CVD, diabetes, hypertension, human immunodeficiency virus (HIV), and hepatitis [4]. Oral health professionals are une-

quivocally capable and willing to identify patients in the early stages of systemic disease and to offer prevention strategies. Furthermore, given that an estimated 19.5 million people in 2008 visited a dental practice and did not seek primary care services, dentists have an existing market for this new model of care [7].

Implications

Conducting a market analysis in addition to drafting a business model development for incorporating limited primary care into the dental practice has several implications. First, proven and sustainable business models could change the future education system and practice of dentistry as the evolution of dentists to oral physicians becomes more tangible. Rather than concentrating on the diagnosis and treatment of oral diseases, more dental programs and residencies would require students to attain a deeper knowledge of chronic, systemic health conditions and how to identify them. Second, integration of oral and systemic health in the dental setting would mitigate the current burden of primary care personnel shortages and improve systemic health out-comes through early diagnoses and subsequent referrals for treatment and lifestyle adjustments. Furthermore, successful implementation would lower overall health care costs and support national economic stability. Lastly, these models could also improve the state of oral health for mil-lions of individuals, as the option of having medical screenings could incentivize patients to seek regular dental care. These implications hinge on a practice model that takes advantage of the unique training of dentists, the availability of dental auxiliaries to handle minor procedures, and growing opportunities for primary care coordination among health professionals.

Methods

Part I. Market Analysis

In order to develop new practice models, it was important to understand the eligible target population of patients: the pool of patients who regularly see a dentist but not a physician (**Figure 1**). After isolating this patient group, we then determined the prevalence of undiagnosed health conditions as measure of the market demand for primary care screening performed by dentists. A market analysis was conducted using health and demographic data from the Center for Disease Control's National Health and Nutrition Examination Survey (NHANES). Sample proportions used to estimate the exact number of susceptible patients were based on the US population. Data abstraction was performed using the SAS Universal Viewer software (SAS Institute Inc.) and Microsoft Excel.

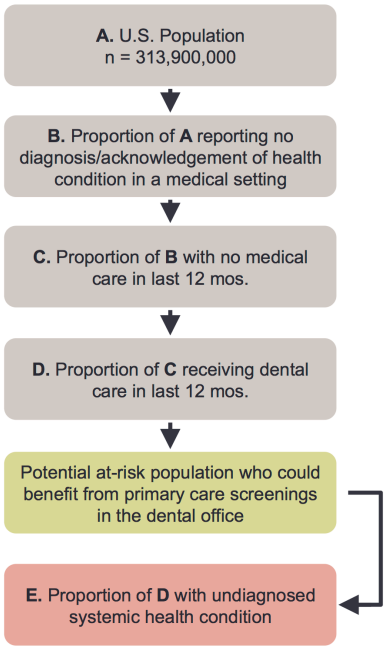


Figure 1. Oral Physician Market Identification Process Flow Chart.

The six leading causes of morbidity and mortality were used to quantify the eligible patient pool. These were diabetes, hypertension, hypercholesterolemia, obesity, tobacco use (smoking), and alcohol use. For each condition, we first determined the appropriate demographic for screening (i.e., women over 45 or adults 20 and over) and extracted the unique patient identifier numbers for that demographic from the 2003–2004 NHANES data. From this population, we then isolated the patients who denied having a diagnosis of, or receiving consultation for, the medical condition of interest. From this patient group, we then isolated those who received no medical care in the previous year by identifying

the respondents who had answered "none" for the number of times they had received health care in past 12 months. "Health care" included visits to the doctor's office, clinic, and/or the hospital emergency room. From this pool, we extracted the patients who received dental care in the last year by finding the respondents who answered "6 months or less" or "not more than 1 year ago" to the question, "How long has it been since you last visited a dentist." This included visits to specialists and/or dental hygienists. Subsequently, we used the NHANES laboratory and medical evaluation data for that year to quantify the number of patients in the sample who did have a previously undiagnosed condition. Since this group received dental but not medical care in the last year, our sample represented the group of patients who could benefit from primary care screening by an oral physician.

Part 2. Costing Analysis

The most recent cost estimates for average life-long treatment associated with the afore-mentioned conditions were identified in the literature. Only direct costs, that is, those solely associated with medical treatments that pertain to the given health conditions, were assessed. These costs were multiplied by the estimated number of individuals with undiagnosed conditions who regularly visit the dentist and thus could benefit from medical screening(s) performed in a dental office. Combined, these costs represent the total economic burden for undiagnosed medical conditions for patients who have seen a dentist but not a physician in the last year. Costs to dentists and the US health care payer system for providing necessary screening and counseling services may be subtracted from the total economic burden in order to calculate the net potential savings that oral physicians can offer in terms of medical expenses averted. These latter costs require more sophisticated analyses and were not part of this work.

Part 3. Business Model Development

Potential models for dentists to address the primary care shortage were theorized based on the following criteria. First, a reasonable model must allow the dentist ample time within one visit to perform dental services as needed in addition to limited primary care screenings. Second, the model should address compensation for the oral physician, and third, the model should provide considerations for referrals, care coordination, and case management for the patients' oral and system conditions. The following assumptions were therefore made in developing the oral physician practice models:

Assumption 1

Traditional health insurance plans could not be used as a source of compensation due to practical constraints. Currently, medical and dental procedure codes are based on two different systems. Cross-coded claims are accepted by some carriers when screening Hb1Ac, CRP, HPV, and HIV, among other tests. Cross coding has had limited adoption by dentists, however, perhaps due to cumbersome medical coding and regulatory issues. Carriers may not allow dentists to screen for certain conditions. Additionally, some states have laws prohibiting dentists from certain screenings, such as HIV testing. Given these concerns, payment models for this new type of practice would have to rely on less conventional means of reimbursement.

Assumption 2

Dentists and their staff are already well qualified and equipped for performing screening services. This is a reasonable assumption, as most, if not all, dental programs, including dental hygiene and assisting schools, educate students on screening principles for various health conditions. Dentists can provide additional education on pathophysiology and management of various health conditions for their staff and patients as needed. Screening tools such as glucose monitoring devices and blood pressure cuffs, if not already in the practice, require little investment and upkeep by the dental office as compared to imaging tools, handpieces, units, and other standard instruments of dental operatory.

Assumption 3

Oral physicians and their auxiliaries will be able to successfully recognize the specific systemic health

conditions from our market analysis and will be capable of making appropriate referrals to medical general practitioners and specialists to help manage these patients.

Results

Part 1. Market Analysis

Diabetes

The market for diabetic patients was identified by those reporting no previous diagnosis of diabetes, in addition to having received dental care but no medical care within the last 12 months (**Table 1**). Pre-diabetes and diabetes were identified as having an HbA1c of 5.7 or higher.

Table 1. Diabetic Screening Market.

Population	N	Proportion	Pop. Estimate
Adults 20 and over	5041	0.498	156,329,767
…No reported diabetes	4425	0.878	137,226,585
…No medical care	699	0.158	21,677,148
…Received dental care	231	0.330	7,163,693
…Undiagnosed diabetes	2	0.009	62,023
…Undiagnosed pre-diabetes	23	0.101	721,020
…Total undiagnosed diabetes and pre-diabetes	25	0.117	783,043

Estimates for US population who could benefit from diabetes screening by a dentist. The initial sample of NHANES respondents was 10,122 individuals, which was equated with the total US population of 313,900,000 at the time of study. Population estimates were determined by multiplying successive proportions from the previous sample by the previous population estimate. Similar algorithms were used to assess markets for other systemic health conditions.

For every dentist, there are five patients with undiagnosed pre-diabetes or diabetes. According to our analysis, there were approximately 7 million people in the U.S above the age of 20 who reported no diagnosis of diabetes, did not seek medical care in the past year, but did see a dentist in the past year. Of this group, 11.7%, or 783,043, were pre-diabetic or diabetic according to lab studies conducted by NHANES.

Hypertension

The United States Preventive Services Task Force (USPSTF) recommends blood pressure screenings for all adults 18 and over. Hypertension, defined as having a systolic blood pressure (SBP) of 140 mm Hg or higher or diastolic blood pressure of 90 mm Hg or higher, is a key risk factor for stroke, heart attack, and heart failure. Additional complications stemming from high blood pressure include aneurysm of arteries, retinopathy, and end-stage renal disease.

To calculate the market for hypertensive patients, we included all men and women in the NHANES study over age 18 with the following criteria: (1) absence of medical history of hypertension, (2) did not see a physician (did not receive health care services) in the past 12 months, and (3) did see a dentist in the past month (**Table 2**). From this sample, we identified the group of patients who had undiagnosed hypertension as defined by the USPSTF criteria outlined above. Only individuals were with an average systolic pressure ≥140 mm Hg or diastolic pressure ≥90 were included in the target sample.

Table 2. Hypertensive Screening Market.

Blood Pressure	N	Proportion	Pop. Estimate
Only systolic ≥ 140 mm Hg	9	0.040	279,115
Only diastolic ≥ 90 mm Hg	5	0.022	155,086
BP ≥ 140/90 mm Hg	9	0	0
Total	14	0.062	434,241

US population estimates of those eligible for hypertension screening by a dentist.

Our estimate for the total population who would benefit from hypertension screening by a dentist was 434,000, which includes patients whose systolic or diastolic blood pressures were in the hypertensive range.

Hypercholesterolemia

High levels of LDL cholesterol can lead to atherosclerosis, which leads to the deposition of plaques in arteries. Progression of atherosclerosis in turn

can lead to arterial stenosis or plaque rupture, both of which can culminate in complete arterial occlusion leading to myocardial infarction (heart attack) or stroke.

USPTF guidelines recommend lipid screening, including cholesterol measurement for all adults males ages 35 and above and all women ages 45 and above. Samples were therefore stratified according to Task Force guidelines.

A summary of results is presented in **Table 3**. Upon analysis of the laboratory data for men over 35, there were no men from this group who had undiagnosed hypercholesterolemia in terms of LDL level (LDL >160 mg/dL). While none of the members of the target group had high cholesterol (total or LDL), approximately 10% had high total cholesterol with borderline high LDL levels (130–159 mg/dL). The latter figure can be extrapolated to 67,370 males over 35 who have undiagnosed borderline high LDL cholesterol who may benefit from a lipid panel performed by a dentist.

When examining the population of women over 45, we found a sample of 21 people from 1,466 who had (1) not seen a doctor in the last year, but (2) had seen a dentist, yet (3) had never been diagnosed with high cholesterol. When extrapolated, this represents approximately 651,000 women over 45 who could potentially benefit from a lipid screening by a dentist. Upon analysis of the examination data, there were 3 of the 29 (14%) with either borderline high (130–159 mg/dL) or high (>160 mg/dL) LDL cholesterol levels. This can be extrapolated to 93,035 women over 45 with undiagnosed borderline high or high LDL cholesterol who could benefit from a lipid screening performed by a dentist.

In sum, this analysis estimates that 160,405 adults of the recommended age for cholesterol screening (men 35+, women 45+) have undiagnosed borderline-high to high LDL cholesterol. Upon including those who may have known about their cholesterol level(s) but see a dentist more often than a doctor, there were an additional 552,862 with borderline high to high LDL levels. Thus, dentists can potentially diagnose and/or monitor lipid levels for an estimated 713,267 men and women in the US.

Table 3. High-LDL Cholesterol Screening Market.

LDL Cholesterol	N (M)	N (W)	Men over 35	Women over 45	Pop. Estimate
High (>160 mg/dL)	0	1	0	31,011	31,011
Borderline High (130-159 mg/dL)	1	2	67,370	62,023	129,939
Total	1	3	67,370	93,035	160,405

US population estimates of those with undiagnosed hypercholesterolemia who visited a dentist but not a doctor in the past 12 months.

Obesity

A summary of results is presented in **Table 4**. Our estimate of all obese individuals (obesity defined as BMI greater than or equal to 30 kg/m^2) receiving regular dental but no medical care was approximately 2.2 million. From this population, approximately 28% had never been told by a physician that they were overweight. Therefore, we conjecture that approximately 630,000 individuals could benefit from obesity screening and counseling by a dentist. A key assumption here is that a physician not telling these individuals that they are overweight also implies that they have never counseled their patients on weight loss.

Table 4. Obesity Screening Market.

Population	N	Proportion	Pop. Estimate
Over age 2	9644	0.919	288,532,464
…No medical care	1188	0.123	35,542,987
…Received dental care	457	0.387	13,672,681
…& Obese (BMI > 30 kg/m^2)	74	0.162	2,213,957
…Never told overweight by health professional	21	0.284	628,285

US population estimates of those eligible for obesity screening by a dentist.

Tobacco Use

A summary of results is presented in **Table 5**. Because there was no question in the 2003–04 NHANES regarding whether a physician had previously performed tobacco screening or offered tobacco cessation counseling, we used a proxy question:

"Has a doctor ever told you had emphysema?" The question was chosen because emphysema, or chronic obstructive pulmonary disorder, is an illness that is directly caused by smoking (smoking comprises 80% of cases). If individuals are diagnosed with emphysema and are current smokers, it is safe to assume that the physician has offered tobacco cessation advice to those patients. This does provide a limitation in that individuals who may not have smoked for a long enough time would not have COPD, yet they still may have received counseling by a physician sometime before the last 12 months. Therefore, our analysis likely underestimates the number of patients who could benefit from cigarette/tobacco screening performed by a dentist. Nonetheless, this analysis estimates that there are approximately 2.23 million individuals who might benefit from tobacco screening/counseling by a dentist.

Table 5. Tobacco Screening/Counseling Market.

Population	N	Proportion	Pop. Estimate
Adults 20 and over	4980	0.492	154,438,056
…No reported emphysema	4862	0.976	150,778,680
…No medical care	696	0.143	21,584,114
…Received dental care.	234	0.336	7,287,740
…Current smoker	14	0.308	2,232,839

US population estimates of those eligible for smoking cessation counseling by a dentist.

Alcohol Use

Because there were no questions in NHANES involving physician intervention regarding alcohol use, we used the proxy question: "Has a doctor or other health professional ever told you that you had any kind of liver condition?" Our outcome for diagnosing alcohol abuse was an incidence of drinking 5 or more drinks in one day. Using these criteria, we determined that approximately 2.23 million individuals might benefit from alcohol abuse screening by a dentist (**Table 6**).

Table 6. Alcohol Abuse Screening Market.

Population	N	Proportion	Pop. Estimate
Adults 20 and over	5041	0.498	156,329,734
…No reported liver disease	4861	0.964	150,747,199
…Reported drinking 5+ at least once	959	0.197	29,740,161
…No medical care	224	0.233	6,946,607
…Received dental care.	72	0.321	2,232,848

US population estimates of those eligible for alcohol screening by a dentist.

Market Analysis Summary

Figure 2 shows the total populations representing the market for dentist-led primary care screenings for alcohol, tobacco, obesity, high cholesterol, hypertension, and diabetes. The three largest markets for medical screenings performed by oral physicians are in alcohol- and tobacco-using populations, as well as the diabetic/pre-diabetic population, which is approximately ~65% smaller. The smallest screening market for dentists is in the high cholesterol population.

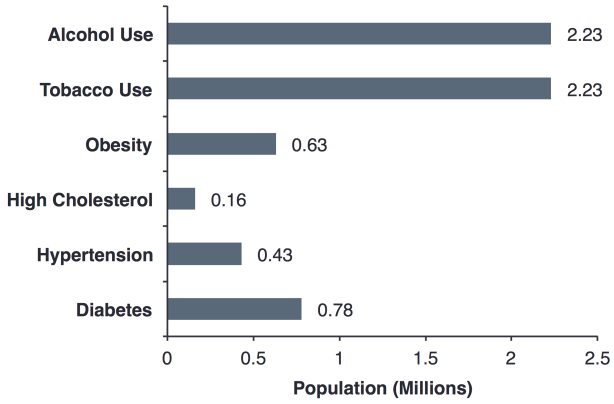


Figure 2. Population with Undiagnosed Conditions that May Benefit from Dentist Screening.

Part 2. Costing Analysis

Summaries of the costing analysis are presented in **Table 7** and **Figure 3**. Our costing analysis represents the consequence of dentists failing to perform primary care screenings in their offices. The strength of this assessment is that it uses the most up-to-date economic data regarding the direct medical costs of unaddressed systemic health condi-

tions. Literature searches were conducted to determine a cost per patient in direct medical expenditures [8–12]. This cost was then multiplied by the market population estimate for each condition to determine the total economic burden for each condition.

Table 7. Costing Analysis Calculation.

Condition	Pop. Estimate	Yearly Medical Costs/Case	Total Costs
Diabetes	783,043	$8,294	$6,494,558,642
Hypertension	432,241	$2,781	$1,202,062,221
Hypercholesterolemia …Undiagnosed	160,405	$158	$25,263,788
…Pre-diagnosed	552,862	$158	$87,075,765
Obesity	628,285	$2,741	$1,722,129,185
Tobacco Use	2,232,839	$2,192	$4,894,383,088
Alcohol Use	2,232,848	$348	$777,031,104
Total			$15,202,503,793

Sample calculation: 783,043 N × $8,294 DME = $6,494,558,642 TEB, where N = size of the target market, DME = direct per patient medical expenditures for the undiagnosed (and untreated) condition, and TEB = total economic burden from the undiagnosed condition

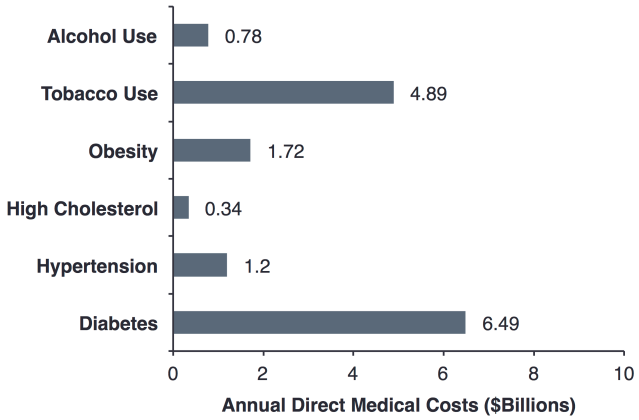


Figure 3. Total Burden of Undiagnosed Illnesses in Target Population.

Part 3. Business Model Development

Based on the assumptions presented in Part III of the Methods section, we devised two practice models that can allow the oral physician to provide quality oral health care in addition to primary care screening services: a concierge practice model and a "partner-practice" model contracted with an Accountable Care Organization.

The first model is a concierge practice. Concierge medical practices, as defined by the Annals of Internal Medicine, provide "expanded access to care and individualized attention, collect charges from insurance companies and directly from patients. Some bill hundreds of dollars for one-time 'executive' physicals, whereas others have patients pay annual retainer fees" [13]. A concierge practice fits well with the roles of an oral physician. Patient subscriptions will be able to cover a myriad of benefits, from immediate phone/email access to the dentist to next-day appointments and complimentary services such as teeth whitening. Patients can still use their dental insurance for certain procedures. Best of all, the subscription will compensate the dentist for primary care screening activities. Increased fees generated by the concierge practice can be balanced by a reduction in the dentists' patient pool. This will allow the oral physician to allocate ample time for providing quality dental treatments in conjunction with medical screenings. A schematic for this model is presented in **Figure 4**.

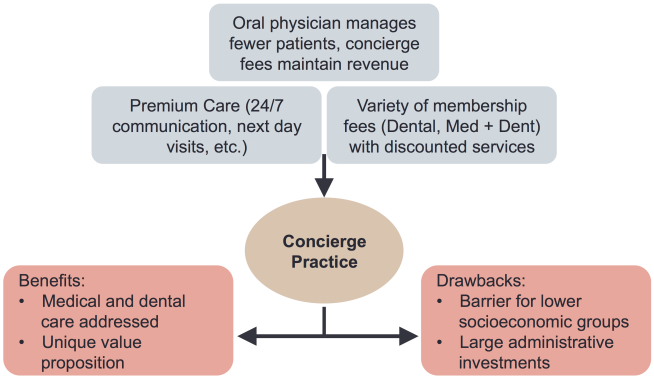


Figure 4. The Concierge Model for the Oral Physician.

This concierge practice model is advantageous for patients desiring a premium level of dental care with the added benefit of general health screening. This type of practice may, however, exclude patients from lower socioeconomic backgrounds depending on the fee structure of the practice. Transitioning an existing dental practice into this oral physician model may also be difficult if patients must be dropped from the practice to decrease the size of the patient pool. Nonetheless, the concierge care

model represents an innovative and attractive business offering for newly minted dentists to address the primary care conundrum of our generation.

The second model we developed can be best described as a "partner-practice" contracted with an Accountable Care Organization and is presented in **Figure 5**.

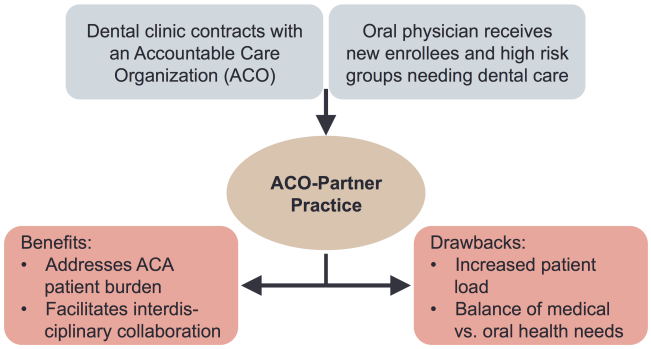


Figure 5. The ACO Contract Model for the Oral Physician.

An Accountable Care Organization, or ACO, is a group of health care providers and institutions that is responsible for managing the health of a specific subset of patients. According to Marko Vujicic of the American Dental Association, "two essential features of the ACO are 1) designated accountable provider entities which share responsibility for treating a group of patients, and 2) performance measurement and new reimbursement mechanisms" [14]. Oral health policy experts agree that dental care integration into ACOs would be an important step towards coordinating oral and systemic health care in an integrated effort while reducing health care costs. Despite this sentiment, dental-ACO arrangements are being challenged by the fact that dental care is not considered a core service in need of integration [14]. Adult dental care is also not an essential health benefit under the insurance plans of most ACO patients.

Our model addresses these issues by allowing dentists to play a larger role in the overall health of an ACO patient pool. As ACOs move towards global per-patient budgets, dental offices that contract with an ACO can receive a portion of that budget for performing primary care screenings along with bonuses for correct diagnoses. Since the ACO itself would implement the entire infrastructure surrounding reimbursement and referrals, the oral physician would need little capital to adjust the practice model. Through shared financial risk, co-location, and electronic medical record technology, oral physicians could become part of a health care team that effectively coordinates oral and general health leading to reduced overall costs incurred by public insurance [14].

A potential drawback to the "partner-practice" model is that dental practice groups may have little influence in the coordination effort depending on how the ACO leadership views dental-medical integration. In addition, because ACOs are designed to address high-risk groups in the Medicaid and Medicare population, dental offices may not be able to keep up with the demand along with reduced reimbursement from government insurance programs.

Discussion

Part 1. Market Analysis

Given the available data, our analysis represents the most accurate prediction of the market that can be targeted by dentists for primary care screening. However, several key limitations exist within this methodology. Foremost, the NHANES demographic and examination data were obtained in 2004. While more recent medical examination data were available, the 2003–2004 NHANES is the most recent survey that includes questions on dental care utilization, without which we would have been unable to identify our target market. It is likely that an analysis with current data on health care utilization and population health would yield different results, given the greater access to primary care offered to millions by the Affordable Care Act and potential changes in dental care utilization over the past decade.

One factor that confounds our analysis is that several data points in the NHANES survey are missing. For example, of the total 10,122 individuals in the survey, only 4,034 were tested for LDL-cholesterol,

of which 416 individuals had missing values for their cholesterol data. In a sample of 6,213 individuals who were asked if their doctor had informed them of their high cholesterol level, 2,727 (44%) had no recorded response. That fact that such a large portion of our sample could not be measured suggests that our analysis most likely underestimates the market for cholesterol screening by dentists. A similar phenomenon occurred in the tobacco market assessment. In a sample of 5,041 people, there were 2,541 missing responses regarding smoking habits. Furthermore, when assessing the number of patients reporting a diagnosis of emphysema, there were 5,040 missing responses (52%) out of a possible 9,645. The extent of these missing data suggests a significant underestimation of the available market for tobacco screening by dentists. Missing data points also impinged upon the obesity market analysis. In a sample of 9,645 respondents, 4,605 individuals (47.7%) had no recorded response for the question, "Has a doctor said you were overweight?" Therefore, the population that stands to benefit from obesity screening/counseling by a dentist could be much larger than 630,000.

The market for primary care screening by dentists elucidated by this analysis is much smaller than it might be in practice. While these population estimates refer to individuals who are unaware of their condition, there is still potential for dentists to screen for individuals who have already received a diagnosis. A patient could have had high cholesterol in the past and might need continuous monitoring by a dentist is he/she regularly does not seek medical care.

Part 2. Costing Analysis

The costing analysis shows that dentists offering limited primary care services could save the health care market approximately $15.2 billion per year. It is important to note that the costing studies in our analysis each used different methodologies, and therefore the costs per patient per year relating direct medical expenditures may be variable according to disparate estimates for various medical procedures. Furthermore, one costing study was published in 2009 and another in 2000, which may alter the true medical costs of neglecting tobacco and alcohol abuse, respectively. Certain medical expenditures could have decreased due to technological advances or increased due to inflation.

In addition, it may not be possible for the dental profession to completely eliminate these diseases based on their diagnoses. First, dentists may not be able to screen for such conditions with perfect accuracy. Additionally, medical interventions are never completely successful. Lifestyle adjustments may take years to become fully ingrained, and medication efficacy is always determined by patient compliance. Furthermore, a dentist's counseling may not be as strong an influence towards a healthier lifestyle as that of a medical physician.

Part 3. Business Model Development

The business two models presented above represent possible directions in which the oral physician can take if he/she endorses a comprehensive model of treatment focused on quality dental care along with limited primary care screening appropriate for the dentist's training. We do not assert these are the only two models that can be best utilized by the oral physician. In fact, comprehensive care payment models, a blend of capitation (global budgets) and pay-for-performance measures, may be a better approach because they base a significant portion of the practice's income on achieving valued outcomes and specific performance goals [15]. Risk-adjusted payment models will incentivize care for those who need it most, while investments in technology and team-based care can improve the overall patient experience. Regardless of which practice model is ideal, the business models of the oral physician must be evaluated in a number of practice settings in order to address panel size, training, staffing considerations, scheduling and more. Insurance corporations and the Centers for Medicare and Medicaid Services should carry out these studies and disseminate their findings to dental schools, dental service organizations (DSOs), and dental professional organizations.

Conclusions

The purpose of this work was to identify the target market for the oral physician to address the United States primary care crisis. Using the most recently available national health statistics on dental and medical care access, including physical and laboratory examination data, we were able to identify a large segment of the population with undiagnosed health conditions, including diabetes, hypertension, hypercholesterolemia, obesity, tobacco use, and alcohol use. An estimated 6.5 million people have one of these conditions and are unaware of it. Because these people do not regularly seek medical care, they are vulnerable to progression and acute exacerbations of these conditions. This inevitably leads to increased mortality and increased medical expenditures, which largely come from the taxpayers' pockets. However, many of these people who fall between the healthcare cracks do regularly visit dentists, placing dentists in a strategic position to address these gaps in care. Therefore, dentists must be able to redefine themselves as oral physicians and assume their responsibility for performing quality dental care in conjunction with limited primary care screenings.

The business of dentistry has been static for the past hundred years. Dentistry must adapt towards new and innovative delivery and financing models for the evolution of dentists into oral physicians. Prospective models include a concierge practice, which charges patients a retainer fee for holistic care, and an ACO-contracted practice that facilitates seamless referrals and care coordination with other medical professionals. If dentists and dental specialists are able to adapt to the changing climate of health care and become more involved in the systemic health of their patients, they will be able to save millions of lives and billions of dollars in avoided medical expenditures. Furthermore, dental practice innovation will undoubtedly lead to increased research efforts on the relationship between oral and systemic disease. It could also inspire more of our nation's best and brightest minds to join the dental profession and make greater advances toward a healthier America.

References

1. Green LV, Savin S, Lu Y (2013) Primary Care Physician Shortages Could Be Eliminated Through Use of Teams, Nonphysician, and Electronic Communication. *Health Affairs* 32(1):11-19.
2. Bodenheimer T and Hoangmai HP (2010) Primary Care: Current Problems and Proposed Solutions. *Health Affairs* 29(5): 799-805.
3. Greenberg BL, Glick M, Frantsv-Hawley J, Kantor ML (2010) Dentists Attitudes Towards Chairside Screening for Medical Conditions. *J Am Dent Assoc* 141:52-62.
4. Group O, Angeles L (2008) New Opportunities for Dentistry in Diagnosis and Primary Health Care. *J Dent Educ* 72(2 Supplement):66-72.
5. Giddon DB (2006) Why Dentists Should Be Called Oral Physicians Now. *J Dent Educ* 70(2): 111-114.
6. Giddon DB (2012) Oral Physicians. *Brit Dent J* 213(10): 497-498.
7. Giddon DB (1999) Mental Dental Interface: Window to the Psyche and Soma. *Perspect Biol Med* 43(1): 84-97.
8. (2010) The United States of Diabetes: Challenges and Opportunities in the Decade Ahead. *United Health Center for Health Reform & Modernization*.
9. Heidenreich PA, Trogdon JG, Khavjou OA et al. (2011) Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association. *Circulation* 123(8): 933-44.
10. Cawley J, Meyerhoefer C (2012) The medical care costs of obesity: an instrumental variables approach. *J Health Econ* 31(1): 219-30.
11. Kahende JW, Adhikari B, Maurice E, Rock V, Malarcher A (2009) Disparities in health care utilization by smoking status—NHANES 1999-2004. *Int J Environ Res Public Health* 6(3): 1095-106.
12. Fleming MF, Mundt MP, French MT, Manwell LB, Stauffacher EA, Barry KL (2000) Benefit-cost analysis of brief physician advice with problem drinkers in primary care settings. *Med Care* 38(1): 7-18.
13. Stillman, M (2010) Concierge Medicine: A Regular Physician's Perspective. *Ann Intern Med* 152(6): 391-392
14. Vujicic, M and Kamyar N (2013) Accountable Care Organizations Present Key Opportunities for the Dental Profession. *American Dental Association Health Policy Institute*.
15. Goroll AH, Berenson RA, Schoenbaum SC, Gardner LB (2007) Fundamental reform of payment for adult primary care: comprehensive payment for comprehensive care. *J Gen Intern Med* 22(3): 410-5.

HMSR

REVIEW
Clinical

Preventing Stroke in Atrial Fibrillation: Warfarin vs. Factor Xa Inhibitors

Pratyaksh K. Srivastava, BA[1]

[1]Harvard Medical School, Boston, MA 02115, USA
Correspondence should be addressed to P.K.S. (Pratyaksh_Srivastava@hms.harvard.edu)

Warfarin has traditionally been used to prevent stroke in patients with atrial fibrillation. The advent of factor Xa inhibitors, however, has provided an alternative. This article evaluates the literature on factor Xa inhibitor use in patients with atrial fibrillation and discusses the general benefits and risks of this novel class of anticoagulants.

Introduction

The heart is a muscular pump that circulates blood to the body. It accomplishes this through an electrical network that dictates the timing and nature of its contraction. Atrial fibrillation is an electrical abnormality whereby the heart beats irregularly and too quickly [1]. One consequence of atrial fibrillation is an increased tendency to form blood clots in the upper chambers of the heart [1]. These blood clots have the ability to break off and travel to the brain, where they can cause a stroke.

Patients with atrial fibrillation are at five times greater risk for stroke than similar aged individuals in normal sinus rhythm [2]. In order to prevent strokes, patients with atrial fibrillation are frequently anticoagulated with warfarin, a Vitamin K antagonist.

The decision to begin anticoagulation is usually based on a patient's $CHADS_2$ (history of congestive heart failure and/or moderate or severe systolic dysfunction; hypertension; age ≥ 75; diabetes mellitus; prior stroke, transient ischemic attack, or thromboembolism) score (**Table 1**) [3,4]. US guidelines recommend the initiation of warfarin therapy in patients with $CHADS_2$ scores of 2 or above, which correlates to an ischemic stroke rate of 4.2% per year (**Table 2**) [3,4].

Table 1. Calculation of $CHADS_2$ Score.

$CHADS_2$	Score
C: Congestive heart failure	1
H: Hypertension	1
A: Age ≥ 75	1
D: Diabetes mellitus	1
S: Stroke, transient ischemic attack, or thromboembolism	2

Table 2. Ischemic Stroke Rate (%/year) by $CHADS_2$ Score.

$CHADS_2$ Score	Ischemic Stroke Rate (%/year)
0	0.6%
1	3.0%
2	4.2%
3	7.1%
4	11.1%
5	12.5%
6	13.0%

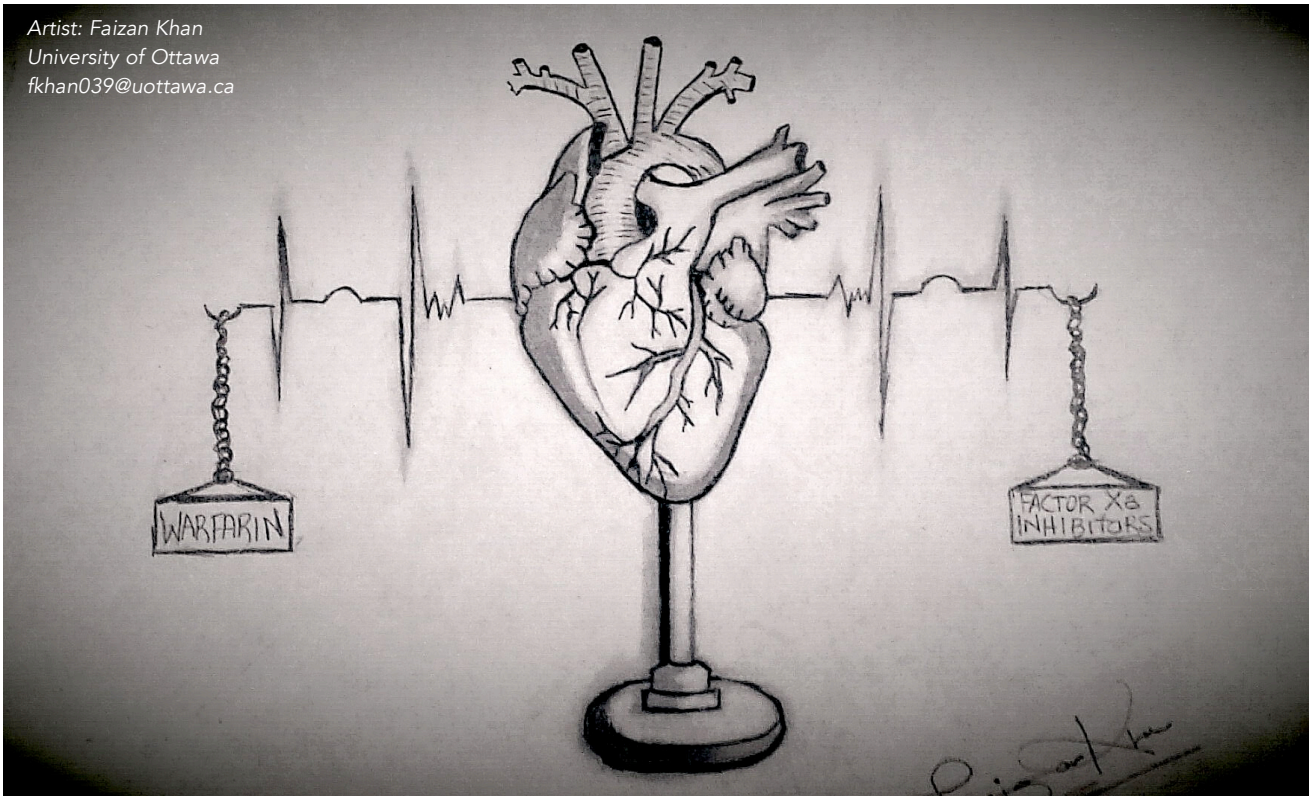
Currently, only 60% of patients with atrial fibrillation who meet criteria for anticoagulation take warfarin [5]. The main reasons cited for this low rate include fear of bleeding complications, difficulty dosing the medication, and the bothersome nature of frequent blood draws [5]. These associated difficulties were the impetus for trialing direct Xa inhibitors, which anticoagulate by blocking factor Xa in the coagulation cascade. Xa inhibitors were initially attractive for their ease of dosing (few drug and food interactions, standard dosing regimens) and for their convenience (no frequent blood draws) [5].

CLINICAL TRIALS: XA INHIBITORS VS. WARFARIN

The first two large randomized controlled trials (RCTs) to compare Xa inhibitors to warfarin were "Rivaroxaban versus Warfarin in Nonvalvular Atrial Fibrillation" (ROCKET AF) and "Apixiban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation" (ARISTOTLE) [6,7].

ROCKET AF was a randomized, double blind, controlled trial performed between 2006 and 2010 that compared oral rivaroxaban to warfarin for the prevention of stroke or systemic embolism in atrial fibrillation patients [6]. The trial enrolled 1,178 individuals with atrial fibrillation and moderate to high risk of stroke ($CHADS_2$ ≥2). The patients were randomized to receive once daily oral rivaroxaban or dose-adjusted warfarin. Results of the ROCKET AF study revealed a decreased incidence of stroke and systemic embolism in the rivaroxaban group when compared to the warfarin group (HR 0.79, CI 0.66–0.96). The rivaroxaban group also demonstrated a significant reduction in intracranial hemorrhage (0.5% vs. 0.7%, p=0.02) and fatal bleeding (0.2% vs. 0.5%, p=0.003) when compared to the warfarin group.

The ARISTOTLE trial (2006–2010, n=18,201) used the Xa inhibitor apixiban and was similar to ROCKET AF in design and outcomes measured [7]. ARISTOTLE demonstrated similar findings to those of ROCKET AF, with a reduced incidence of stroke and systemic embolism in the apixiban group when compared to the warfarin group (HR 0.79, CI 0.66–0.95), as well as reductions in major bleeding (HR 0.69, CI 0.60–0.80), hemorrhagic stroke (HR 0.51,



CI 0.35–0.75), and all-cause mortality (HR 0.89, CI 0.80–0.99). Neither study reported major side effects of the drugs beyond the bleeding described above.

One limitation of ROCKET AF was that the warfarin group had a therapeutic International Normalized Ratio (INR) only 55% of the time. Previous studies have demonstrated that a therapeutic range more than 58% of the time is required for warfarin anticoagulation to be considered superior to aspirin [8]. There was also a significant difference in number of prior myocardial infarctions between the rivaroxaban and warfarin groups [6]. The warfarin group exhibited a greater number of prior myocardial infarctions, suggesting the possibility of greater underlying cardiovascular disease burden. Both studies evaluated individuals at moderate-high risk for stroke—mean $CHADS_2$ for ROCKET AF and ARISTOTLE were 2.1 and 3.5, respectively—and so may not generalize to individuals with lower $CHADS_2$ scores. Pharmaceutical companies sponsored both ROCKET AF and ARISTOTLE, offering a potential source for bias in result interpretation.

Many other Xa inhibitors have been evaluated since ROCKET AF and ARISTOTLE. In 2013, a Cochrane Review was released pooling the major RCTs on the topic [9]. The review included 43,084 participants from ten RCTs that compared dose-adjusted warfarin to different Xa inhibitors, including apixiban, betrixaban, darexaban, endoxaban, idraparinux, and rivaroxaban for the prevention of stroke and other embolic events in atrial fibrillation patients. The mean $CHADS_2$ score of the study participants was 2.7 (range: 1.9–3.5), and the mean age ranged from 65 to 74 years. In the Cochrane group's meta-analysis, it was found that Xa inhibitors significantly decreased stroke (OR 0.78, CI 0.69–0.89) and systemic embolic events (OR 0.53, CI 0.32–0.87) when compared to warfarin. They were also associated with fewer intra-cranial hemorrhages (OR 0.51, CI 0.41–0.64) and lower all-cause mortality (OR 0.88, CI .81–0.97).

LIMITATIONS OF XA INHIBITOR ADOPTION

There are several limitations to the studies presented above. No studies to date compare the different Xa inhibitors directly against one another (e.g., apixiban vs. rivaroxaban), making it difficult to conclude if one Xa inhibitor is better than another [9]. Further, the majority of the large studies performed to date use either apixiban or rivaroxaban—85% of the data from the Cochrane Review was from studies that used either one of these two drugs. In fact, in sub-group analysis, the Cochrane Review found no significant reduction in stroke or systemic embolic events in any of the Xa inhibitors except apixiban and rivaroxaban. The longest follow-up time for studies included in the Cochrane analysis was 1.9 years, with some studies having follow-up of as little as 12 weeks. Finally, the studies conducted to date all include patients who both have and have not used warfarin before study enrollment, making it difficult to evaluate if prior warfarin status affects Xa inhibitor efficacy [9].

This last limitation was addressed in the "Endoxaban versus Warfarin in Patients with Atrial Fibrillation" (ENGAGE-AF) trial, which was a randomized, double blind trial comparing low- (30mg) and high-dose (60mg) endoxaban to dose-adjusted warfarin for the prevention of stroke or embolic events in patients with atrial fibrillation [10]. The trial enrolled 21,000 patients with atrial fibrillation and moderate to high risk of stroke (mean $CHADS_2$ 2.8) and had a mean follow-up time of 2.8 years. The study found that there was no difference in stroke/embolism rates between the high-dose endoxaban and warfarin groups in individuals with prior Vitamin K antagonist use (1.62% vs. 1.60%, $p>0.05$), while there was a difference in Vitamin K antagonist naïve individuals (1.49% vs. 2.12%, $p=0.03$). These results suggest that prior Vitamin K antagonist status may play a role in determining response to Xa inhibitors.

The clinical adoption of Xa inhibitors has been slow [11]. This can be attributed to a number of factors. Although the new Xa inhibitors are associated with a lower incidence of fatal bleeding and intracranial

hemorrhage, bleeding continues to remain a feared complication due to the lack of effective Xa inhibitor reversal agents. This is especially worrisome in certain populations, such as the elderly or those with gastrointestinal ulcers, where bleeding is particularly dangerous. The current mainstay of treatment for patients who present with acute bleeding on a Xa inhibitor is supportive care with fluid resuscitation, transfusions, source identification, and drug discontinuation [12]. The use of 4-factor prothrombin complex concentrate (PCC) was shown to normalize prothrombin time in human volunteers who received rivaroxaban, although no studies to date have evaluated the use of PCC in patients with acute bleeding [12]. There are certain additional populations in which factor Xa use is contraindicated. These groups include those with prosthetic heart valves (increases risk of thrombus on the valve), pregnant women (lack of clinical data), patients with renal insufficiency (Xa inhibitors have partial renal excretion), and those at the extremes of weight (affects dosing) [12].

Another issue that has slowed adoption of Xa inhibitors is medication adherence. Warfarin has a half-life of around 40 hr, which allows for patients to miss occasional doses without suffering from thromboembolic complications [12]. Xa inhibitors, in comparison, have a half-life of around 12 hr. This short half-life results in the rapid decline of the drug's antithrombotic effect if doses are missed [12]. Last, factor Xa inhibitors are expensive. Even with insurance, the copayments can be prohibitively high, increasing the likelihood of missed doses, splitting pills, and overall non-adherence [12,13].

CONCLUSIONS

Studies to date demonstrate either superiority or noninferiority of Xa inhibitors to warfarin for the prevention of stroke in patients with atrial fibrillation. There are limitations to the data, however. In addition, there are many downsides to factor Xa inhibitor use including the absence of an effective reversal agent, rapid decline of antithrombotic effect if doses are missed, and high cost. Taken together, these findings should prompt caution when considering the use of Xa inhibitors for the prevention of stroke in patients with atrial fibrillation.

REFERENCES

1. Waktare JE (2002) Atrial fibrillation. *Circulation* 106:14-16.
2. Wolf PA, Abbott RD, Kannel WB (1991) Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 22:983-8.
3. Singer DE, Albers GW, Dalen JE, *et al*. (2008) Antithrombotic therapy in atrial fibrillation: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest* 133:546S-92S.
4. Friberg L, Rosenqvist M, Lip GY (2012) Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: The Swedish atrial fibrillation cohort study. *Eur Heart J* 33:1500-1510.
5. Veltkamp R, Prabhakaran S (2014) Critique of effective anticoagulation with Factor Xa Next Generation in Atrial Fibrillation trial. *Stroke* 45:2151-3.
6. Patel MR, Mahaffey KW, Garg J, et al. (2011) Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med* 365:883-91.
7. Granger CB, Alexander JH, McMurray JJ, *et al*. (2011) Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 365:981-92.
8. Connolly SJ, Pogue J, Eikelboom J, et al. (2008) Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. *Circulation* 118:2029-37.
9. Bruins Slot KM, Berge E (2013) Factor Xa inhibitors versus vitamin K antagonists for preventing cerebral or systemic embolism in patients with atrial fibrillation. *Cochrane Database Syst Rev* 8:CD008980.
10. Giugliano RP, Ruff CT, Braunwald E, et al. (2013) Edoxaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 369:2093-104.
11. Bauer KA. Pros and cons of new oral anticoagulants (2013) *Hematology Am Soc Hematol Educ Program* 2013:464-470.
12. Kaatz S, Kouides PA, Garcia DA, et al. (2012) Guidance on the emergent reversal of oral thrombin and factor Xa inhibitors. *Am J Hematol* 87 Suppl 1:S141-5.
13. Choudhry NK, Saya UY, Shrank WH, et al. (2012) Cost-related medication underuse: prevalence among hospitalized managed care patients. *J Hosp Med* 7(2):104-109.

HMSR

REVIEW
Historical

Historical Path of Discovery of Viral Hepatitis

David T. Wong[1]; Martin C. Mihm, MD[2]; James L. Boyer, MD[3]; and Dhanpat Jain, MD[4]

[1]Dartmouth College, Hanover, NH 03755, USA
[2]Harvard Medical School and Brigham and Women's Hospital, Department of Dermatology, Boston, MA 02115, USA
[3]Yale University School of Medicine, Department of Internal Medicine, Section of Digestive Diseases, New Haven, CT 06519, USA
[4]Yale University School of Medicine, Department of Pathology, New Haven, CT 06519, USA
Correspondence should be addressed to M.M. (mmihm@partners.org)

This article presents an overview of the historical timeline of the gradual discovery of various hepatitis viruses and the pivotal roles of epidemiological observations, human experimentations, and laboratory research in their discovery and containment.

Abstract

Viral hepatitis is an ongoing global infectious public health problem and a major cause of chronic liver diseases, including liver cancer. Previously described as "epidemic jaundice," viral hepatitis has been known to exist since ancient civilizations. The contagious nature of the illness was suspected even in the eighth century CE. Records from major military campaigns in different continents from the 18th to 20th centuries, including the American Civil War and the First and Second World Wars, reported that "campaign jaundice" caused significant morbidity of the troops and impacted war strategies. Epidemiological observations from late 19th century and research, including human experimentation in the 20th century, led to the gradual identification of a distinct "infectious hepatitis" agent transmitted by oral-fecal transmission, known later as hepatitis A virus (HAV), and a "serum hepatitis" agent transmitted by inoculation or transfusion of serum, blood or plasma, or sexual contact. Experiments that involved feeding and injecting infected feces, urine, and serum into volunteered military personnel, prisoners, and mentally retarded children raised issues of informed consent and mental competency of the retarded children. Only until the 1960s was one of the causative agents of "serum hepatitis," the hepatitis B virus (HBV), discovered. Further research led to the discovery of additional hepatitis viruses (HCV, HDV, HEV, and HGV). Breakthroughs in the containment of the hepatitis epidemic included development of hepatitis vaccines and recent therapeutic successes for hepatitis C. This paper presents an overview of the historical timeline of the gradual discovery of the causative agents of viral hepatitis.

Introduction

Viral hepatitis refers to inflammation of the liver caused by viral infection. Most cases of viral hepatitis are caused by HAV, the causative agent of viral hepatitis A (previously called infectious hepatitis, infectious jaundice, or campaign jaundice); hepatitis B virus (HBV), the causative agent of viral hepatitis B (previously called serum hepatitis); hepatitis C virus (HCV), the causative agent of viral hepatitis C (previously called Non-A, Non-B hepatitis); or hepatitis E virus (HEV), the causative agent of enter-

ically transmitted hepatitis.

According to the World Health Organization (WHO), viral hepatitis is one of the most common infectious diseases and a global public health problem. About 1 in 12, or 500 million people, are chronically infected with HBV or HCV, with millions more at risk. One million deaths each year can be attributed to viral hepatitis infections and its complications. The majority of the global burden of viral hepatitis is in the Asia-Pacific region.

The path to various discoveries in this field showcases scientific ingenuity but is also full of controversies, especially related to human experimentation. Modern medicine has come a long way since the earliest documentation of jaundice as a manifestation of viral hepatitis, but surprisingly, there are no good reviews in the recent literature highlighting this incredible journey. This review traces various discoveries that eventually led to our current understanding of various hepatotropic viruses and viral hepatitis.

ANCIENT PERIOD

Early epidemics of jaundice were reported in the time of Babylonia and ancient China over thousands of years ago [1]. Hippocrates documented an epidemic of jaundice occurring on the island of Thassos in the fifth century BCE [2]. A description of liver disease including jaundice can be found in the Babylonian Talmud of the fifth century BCE as a cause of fever, malaise, lassitude, stomach problems, and sometimes death [3]. The contiguous nature of "jaundice" was first mentioned in a letter from Pope Zacharis to Saint Boniface in 751 CE, in which Pope Zacharis instructed Saint Boniface, Archbishop of Mainz, to hold off serving Holy Communion to persons with jaundice until all the rest had been served and to bury horses infected with the same condition [4]. Throughout history, hepatitis caused frequent pandemics in Europe, ranking only behind cholera and plague [3].

EPIDEMIC JAUNDICE OR INFECTIOUS HEPATITIS

Disease outbreaks called "Epidemic Jaundice," or "Campaign Jaundice" (hepatitis A infection) plagued armies and civilians during medieval wars and impacted militaries for centuries [3]. The French called the disease *jauniesse des camps*; to the Germans it was *Soldatengelbschut*; scientists called it icterus, in reference to a yellow bird from Greek mythology [3]. Similar conditions were described as infectious hepatitis in the United States, infective hepatitis in England, and Botkin's disease in Russia [5]. Epidemics of jaundice occurred in Europe in the 17th, 18th, and 19th centuries, likely due to the growing and increasingly crowded populations.

The Military records from the British Military Hospitals in Germany from 1761 to 1763 documented epidemics of illness among the troops with fever, jaundice, vomiting, diarrhea, abdominal pain, and occasional fatalities [6]. Similar epidemics were reported among British troops in India and may have been caused by "bad water," "change of diet," and "great repletion after long fasting" [7]. Fatal cases of hepatitis following dysentery were reported in Bengal in 1796 [8]. Jaundice also decimated Napoleon's army during his Egyptian campaign in 1798 [3]. Epidemics of jaundice involving a French army in Pavia during the Italian war coincide with a similar epidemic in the town of Pavia [9]. Army surgeons described similar "epidemic jaundice" in almost all parts of the world, including Europe, Asia, and America, caused by small parasites, denominated germs that were "the producers of serious and even fatal forms of hepatic complaint" [10]. An epidemic of jaundice also inflicted the military and civilian populations in Paris during the Franco-Prussian War in 1870 [3]. In retrospect, we feel that these "epidemic jaundice" or "campaign jaundice" were likely the result of HAV, HEV, or other virus outbreaks.

A misconception of hepatitis as "catarrhal jaundice" by Virchow, who misunderstood the cause of jaundice as the consequence of blockage of the common bile duct by mucus plug, delayed the discovery of the true infectious nature of hepatitis [1,11]. In his

lecture series in 1874, Murchison discussed cases of jaundice independent of obstruction of the bile duct [12]. He described "Epidemic Jaundice" in children in Essen in 1772 with high fatality, in Rotherham in 1862 associated with bad drainage and preceded by fatal outbreak of enteric fever, and London in 1846 after a prevalence of extremely hot weather and outbreaks of enteric fever, and fatal cases resembling those of "acute yellow atrophy of the liver" [12]. Other physicians described an epidemic in the island of Martinique that inflicted 30 pregnant women, 20 of whom died after suffering an abortion of premature labor [9]. We believe it is very likely that these fatal cases represent outbreaks of viral hepatitis with fulminant hepatitis from HBV, HCV, HEV, or other viruses.

During the American Civil War, over 71,000 cases of jaundice were reported among both Union and Confederate troops, with outbreaks of sporadic cases and local epidemics of jaundice with fever and diarrhea, largely attributed to the insanitary conditions of the battlefield [13]. However, the mortality from jaundice was low [1]. In Africa, approximately 6,000 cases of jaundice were reported among the British and Dominion troops in the Boer War in South Africa [1].

During WWI, highly contagious epidemics of "infectious jaundice" were reported in British troops in the Middle East Campaign at Gallipoli and Egypt, affecting up to 25% of the units [14]. No parasites or bacteria were identified, and the condition was thought to be more consistent with hepatitis following a systemic infection rather than catarrhal jaundice from plugging of the bile duct, as postulated by Virchow [14]. The possibility of "blood infection" through person-to-person transmission closely associated with dysentery, enterocolitis, or diarrhea was suggested as a cause of "epidemic jaundice of campaign" in Alexandria, Gallipoli, Mudros, Salonika, and Mesopotamia, with involvement of a third or more of some units [15].

During WWII, 200,000 cases of "epidemic jaundice" occurred among US troops alone from 1942 to 1945 [16]. In total, the German army and citizens suffered over 5 million cases of jaundice [1,17]. Major outbreaks often occurred around the Mediterranean shores, including Palestine, Egypt, and Syria [18]. Investigations of fatal cases revealed no evidence of duodenal catarrh or obstruction of the common bile duct by mucus [18]. Interestingly, military officers were more susceptible to the conditions than other ranks [19]. Potential causes for this phenomenon include the officers' congregation in isolated communities and segregation from civilian population with their own lodgings, social clubs, swimming pools, and improvised portable bathing facilities to conserve water [19].

Facing the need to achieve military success during the Second World War, both the US and Britain intensified their research of the understanding of the cause and transmission of the "campaign jaundice" that had significantly impacted the fighting forces.

DISCOVERY OF HEPATITIS A VIRUS, IMMUNOGLOBULIN, AND VACCINE

In 1945, Stokes and Neefe showed that immune globulin (concentrated antibodies obtained from pooled human plasma) provided protection against illness among children at a summer camp who had been exposed to hepatitis A by either preventing or attenuating the hepatitis A infection [20,21]. "Since then, immune globulin has been used widely for post-exposure prophylaxis." Currently, it "remains an effective intervention for preventing the transmission of hepatitis A to family members and other close contacts of patients who have recently become ill" [21].

In 1973, with the use of immune electron microscopy, Feinstone el al. identified a "spherical 27-nanometer particles" in stool obtained from the feces of hepatitis A patients in the acute stage of the disease [22]. The authors concluded that "the particle was serologically specific for this disease, and every hepatitis A patient tested demonstrated a serological response to this antigen," thus suggesting "that it is the etiologic agent of hepatitis A" [22].

In 1996, Hilleman at Merck developed a hepatitis A vaccine with attenuated HAV. With the development of a childhood vaccination for hepatitis A virus infection, the incidence in the United States has decreased significantly, although "the majority of the world's population is still at moderate-to-high risk for hepatitis A virus infection" [23]. In 2001, more than 10,000 cases were reported in the United States, but "the actual number of cases of hepatitis was probably 5 times that reported, and the number of new asymptomatic infections was probably 10 times the number of reported symptomatic cases" [21]. According to Craig, "[one] third of the US population has serologic evidence of previous hepatitis A infection, with a prevalence ranging from 9 percent among children 6 to 11 years of age to 75 percent among persons 70 years of age or older" [21]. Nevertheless, "the rates of hepatitis A infection in the United States have been decreasing gradually during the past several decades" [21]. The decrease is likely due to the "use of hepatitis A vaccine since 1995 in many communities where the rate of infection had been high," combined with "advances in hygiene, including improved water supplies, enhanced sewage disposal, reduced crowding, augmented food safety, and other factors" [21]. In developing countries, where the hepatitis A vaccine is not readily available, "nearly all people have had hepatitis A infection by early adulthood" [21].

SERUM HEPATITIS OR SERUM JAUNDICE

In 2012, Bar-Gal et al. reported the discovery of the full viral genome of ancient HBV (aHBV) extracted from the liver of a 16th century Korean mummy [24]. The authors concluded that "[the] calculated time of most recent common ancestor suggests that the Korean HBV sequence origin dates back at least 3,000 years and possibly as long as 100,000 years" [24]. The proven existence of HBV in ancient times corroborated Blumberg's landmark discovery of Australian antigen in an aborigine who had never received blood transfusion, which was subsequently confirmed to be Hepatitis B surface antigen (HBsAg) [25]. We agree with Dr. Krugman's opinion that "[I]n retrospect, it was obvious that the Australian aborigine [reported by Dr. Blumberg was] a hepatitis B carrier" with HBsAg [5].

Most authors credited Lurmen with reporting the first epidemic of serum hepatitis in 1883 to 1884 in a Bremen shipyard where the workers received vaccine against smallpox derived from human lymph of cases of Vaccinia (cowpox) [26]. 191 of 1,289 workers developed jaundice within 1 to 7 months after receiving the same lot of lymph, whereas 500 workers in the same shipyard vaccinated with a different lot of lymph were unaffected [26]. Lurmen concluded that "Considering the distribution of cases [accordingly], one must take into account the [vaccination] … as the etiological source of icterus epidemic" [3].

Subsequent outbreaks of injection-associated hepatitis were described. In 1908, McDonald discovered that acute yellow atrophy of liver tends to occur in groups and recognized the possibility of a virus as an infectious cause of the disease [27]. He reported an acute, subacute, and chronic phase of yellow atrophy of the liver with cirrhosis in the chronic cases [27]. He also described the development of jaundice in some patients being treated for syphilis [27].

Observations by Stokes in 1920 reported a dramatic increase of 1,000 percent of patients who developed severe jaundice and arthritis after receiving injections of arsphenamin for syphilis at the Mayo Clinic from August 1917 to July 1920 [28]. After ruling out syphilis or arsphenamin as a cause of the jaundice outbreak, the authors suggested a systemic infection of hematogenous source as the possible cause [28]. Similarly, in 1943, McCallum reported a high percentage of jaundice outbreaks in venereal disease centers after being injected with arsphenamine [29]. He suspected that infectious agents were being transmitted from patient to patient by sharing unsterilized and contaminated syringes and needles containing small amounts of infected blood [29].

In 1937, Findlay reported a total of 52 cases of "acute hepatitis" 2 to 7 months after yellow fever immunization of British troops, featured by malaise, loss of appetite, nausea, vomiting, jaundice, dark

urine, pale stool, and weakness [30]. The authors concluded that the post-vaccinal hepatitis could be due to some virus injected with the serum because the vaccine was filtered and treated to be "bacteriologically sterile" [30]. The clinical presentation of post-vaccinal hepatitis was analogous to outbreak of jaundice following antisyphilitic treatment [30]. Similar clinical observations were described in horses with jaundice and acute necrosis of liver approximately 62 to 78 days after inoculation against horse-sickness virus with serum from previously infected horses [30]. In 1938, Propert reported seven cases of hepatitis in children after an injection of convalescent measles serum; three children died of acute necrosis of liver [31].

In 1942, the US Secretary of War reported that 28,585 cases of jaundice had developed among army personnel between January 1 and July 4, apparently from the use of vaccination against yellow fever [32]. The total number of deaths was 62, with a ratio of one death for every 461 cases [32]. In 1944, Turner described an outbreak of hepatitis affecting 4,083 military persons at Camp Polk, Louisiana, from May 1942 to September 1942, after the use of only one lot of yellow fever vaccine [33]. As a result of the high infection rate in Camp Polk, the entire unit was unable to go abroad [3].

Seefe et al. reported a follow-up of the 1942 hepatitis outbreak linked to the receipt of specific lots of yellow-fever vaccine involving approximately 50,000 US Army servicemen [34]. Follow-up of the veterans who had received the implicated vaccine showed 97% of the group of veterans who received the yellow-fever vaccine and developed jaundice tested positive for antibodies to HBV; 76% of the group of veterans who received the yellow-fever vaccine but remained well tested positive for antibodies to HBV; and only one subject had hepatitis B surface antigen, for a carrier state of 0.26% among recipients of the implicated vaccine [34]. The authors concluded that HBV caused the outbreak with about 330,000 military persons possibly infected, and resultant rare occasions of a Hepatitis B carrier state, but induced Hepatitis B antibodies that persisted for life [34].

In 1943, the British Ministry of Health published a memorandum on "Homologous Serum Jaundice" describing an outbreak of jaundice involving 86 of the 266 soldiers inoculated with convalescent plasma from mump patients and 12 cases of jaundice in persons who received transfusion of plasma or whole blood [19]. In addition, based on the reported cases of "Syringe-Transmission Hepatitis," researchers believed that an icterogenic agent was present in apparently normal individuals [19]. Outbreaks of jaundice occurred in patients receiving injections or venepunctures in clinics for various reasons, including arsentherapy in syphilitic clinics, diabetic clinics, sanatoria, and arthritis clinics [19]. Potential causes for these outbreaks included the following: a limited number of available syringes that was unable to meet the high demand, thus requiring the sharing of syringes and needles between patients; the common practice during venipuncture of aspirating small amounts of blood into the syringe to ensure insertion of the needle in the vein; and simply washing the syringes, without sterilizing or boiling them (to avoid potentially breaking them) [19]. The incidence rate of hepatitis in syphilitic clinics that did not sterilize used needles and syringes ranged from 30% to 60% of all patients, whereas there was no jaundice in groups receiving injections by sterilized apparatus [19]. The authors concluded: "Hepatitis was being transmitted in the course of venipuncture and intravenous injections" [19].

Importantly, the authors observed that the "incidence of hepatitis tended to be low in clinics where syringes were sterilized between patients, and high in clinics where the syringes were "merely washed" [19]. In addition, the "[high] infectivity of blood during the long incubation period resulted in continuous contamination of needles and transmission to almost all the patients" [19]. Based on these observations, researchers suspected that an infectious agent carried in human blood was being transmitted

Photograph by R. Frederick Lambert
Harvard School of Dental Medicine
frederick_lambert@hsdm.harvard.edu

from patient to patient by means of syringes and needles that had been inadequately sterilized and was causing the serum hepatitis [19].

At that time, two distinctive clinical forms of hepatitis were recognized based on clinical and epidemiological features. One form of hepatitis had short incubation period, transmitted by oral-fecal route, and occurred as epidemics named "infectious hepatitis" or "campaign jaundice" (now designated as HAV). Another form of hepatitis had a longer incubation period and was transmitted by parental injections or blood transfusion, hence the name "serum hepatitis" (now designated as HBV or HCV). Patients with infectious hepatitis were immune to the same type of hepatitis, and patients with serum hepatitis were not immune to the infectious type. In 1947, MacCallum suggested the names of hepatitis A and hepatitis B for these two forms, respectively [35].

ROLE OF HUMAN EXPERIMENTATIONS

Over the past century, human experimentations played a significant role in the discovery of viral hepatitis and led to many controversies. Many of these experiments that were pivotal in the discoveries of HAV and HBV cannot be conducted today due to ethical considerations. One of the key issues prior to the discovery of HBV by Dr. Blumberg in mid-1960s was the inability of the researchers to isolate and propagate the infectious agents that caused the two main types of hepatitis: that is, infectious hepatitis (HAV) and serum hepatitis (HBV). In the 1940s, humans were the only known susceptible host for hepatitis virus proliferation, and attempts to isolate and propagate the viruses in a laboratory by using cell cultures, guinea pigs, hamsters, rabbits, mice and rats, and even non-human primates failed. A series of experiments in humans established the differences in mode of transmission of HAV (infectious hepatitis) and HBV (serum jaundice), demonstrated protective immunity, and established these as clearly separate disorders.

In 1942, Voegt fed volunteers with duodenal fluid obtained from patients with hepatitis, with one in four patients becoming infected [36]. From 1941 to 1942, Cameron injected serum from jaundiced soldiers with infectious hepatitis from Palestine into soldier volunteers, infecting them with hepatitis [37]. In 1944, MacCallum injected serum and sprayed feces from patients with infectious hepatitis into the nose of volunteers [38].

In 1945, Havens fed volunteers with feces and serum, and he injected serum from patients with infectious hepatitis into volunteers, who became infected with hepatitis [39]. The authors also injected the serum of patients with serum jaundice into volunteers, who also became infected with hepatitis [39]. Accordingly, the authors concluded that hepatitis could be caused by oral feeding of stool and serum and by injection of serum from patients with infectious hepatitis [39].

Neefe et al. fed pooled feces from patients with active serum hepatitis to volunteers [40]. None of the patients showed evidence of hepatitis during a 4- to 6-month period, suggesting that the causative agent either was not present in the feces or was not active when administered by the gastro-intestinal route [40]. By contrast, volunteers fed with pooled feces of patients with infectious hepatitis developed hepatitis within 26 days, confirming the observation of other authors that the causative agent is present in feces of patients with active disease [40].

In 1945, Neefe et al. described an outbreak of infectious hepatitis in a summer camp at Poconos Mountains [41]. Volunteers were inoculated with serum, feces, nasopharyngeal washings, and urine from infected camp patients [41]. Research findings indicated that the causative agent responsible for this epidemic of infectious hepatitis was water borne, excreted in the feces of persons infected with the disease [41]. Anything subject to direct or indirect contamination with feces could provide a potential means of transmission [41].

Subsequent human experimentations demonstrated homologous immunity in serum jaundice (or serum hepatitis) and infectious hepatitis, and protective effect of normal human globulin when administered

during the incubation period of epidemic infectious hepatitis. In 1945, Haven et al. confirmed that volunteers convalescing from experimentally induced infectious hepatitis were resistant to re-inoculation with the same strain of virus 6 to 9 months later [42].

In 1946, Neefe et al. orally fed and parentally injected infected feces and serum into different groups of volunteers [43]. The findings revealed that virus from patients with serum hepatitis (or serum jaundice) is present in pooled plasma and induced hepatitis in normal volunteers after being inoculated parenterally but not orally; feces from patients with serum hepatitis failed to induce hepatitis when administered orally or parenterally (Seitz filtrate) to volunteers; onset of hepatitis in serum hepatitis was relatively insidious; and volunteers who had serum hepatitis after inoculation were resistant to reinfection with serum hepatitis virus but susceptible to infection with virus of infectious hepatitis [43]. For infectious hepatitis, virus-induced active hepatitis developed after 17 to 37 days in 73% of volunteers who were inoculated orally but in only 11% of volunteers who were inoculated parentally [43]. The interval from inoculation with infectious hepatitis virus to the onset of hepatitis did not exceed 37 days [43]. Infectious hepatitis virus was present in the blood and feces of patients with active hepatitis due to oral administration of the virus [43]. Volunteers who developed infectious hepatitis were resistant to reinfection to infectious hepatitis virus but were susceptible to parentally injected serum hepatitis virus [43]. The authors observed that after so-called "catarrhal jaundice," patients were resistant to infectious hepatitis virus but not resistant to serum hepatitis virus for 4 to 10 years [43]. These observations suggest an antigenic similarity of the causative agent of so-called "catarrhal jaundice" and infectious hepatitis, and a difference in antigenic properties of the causative agent of "catarrhal jaundice" and serum hepatitis (or serum jaundice) [43].

In 1951, Stokes et al. conducted human experiments at different institutions, such as institutions for the mentally retarded, prisons, student nurses from orphanage, and state training schools, using subjects such as mentally retarded children and adults, prison inmates, student nurses, "boys and girls," and "adult employees" [44]. The authors concluded that gamma globulin was "highly protective in passive immunization against viral (infectious) hepatitis," even at a dosage of "0.01 ml per pound of body weight" [44]. Stokes showed that "injection of gamma globulin, an antibody-rich distillate of human serum, could modulate the clinical course of infectious hepatitis by means of "'passive' immunity" [44]. He predicted that if hepatitis infection occurred during the period of passive immunity induced by the gamma globulin, the "clinical disease would be mild and long-lasting immunity to future infection might result" and accordingly named this theory "passive-active" immunity [44].

Subsequent studies on the natural history, epidemiology, and transmission of viral hepatitis were conducted from 1956 through 1971 at Willowbrook State School (WSS), a state-funded residential facility for mentally disabled persons in Staten Island, NY, on institutionalized mentally disabled children. Hepatitis was endemic at WSS and Saul Krugman, MD, a pediatrician, was asked to investigate and research for ways to prevent and control its spread at WSS. Krugman's experiments on mentally retarded children raised issues of informed consent and mental competency and generated intense ethical debates among researchers and bioethics scholars [45-47].

DISCOVERY OF HBV

HBV was serendipitously discovered by Baruch Blumberg who was interested in studying the role of blood antigen polymorphisms, inherited differences in specific blood proteins, in genetics of disease susceptibility [5,48]. The discovery of HBV led to a blood-screening campaign that significantly reduced the incidence of post-transfusion hepatitis from HBV [48]. More importantly, this discovery led to the development of a highly effective hepatitis B vaccine, which over the next several decades protected hundreds of millions of people from HBV infection and led to a decrease in the incidence of

HBV-associated hepatocellular carcinoma, thus constituting the first effective vaccine in preventing any cancer [48]. Dr. Blumberg was awarded the Noble prize for his contribution in 1976 [1].

In the 1950s, Blumberg obtained blood samples from native populations from remote areas of the planet in order to investigate their genetic differences and how that might impact their susceptibility to different diseases. He also obtained blood samples from hemophiliac patients who had received blood transfusions from multiple donors and, hence, produced antibodies against antigens from donors. In 1965, Blumberg reported the discovery of the "Australian antigen" in the sera of an Australian aborigine, which reacted immunologically with a panel of sera from hemophiliac patients by forming a precipitin line in agar gel immunodiffusion [25]. He speculated that the "antigen" is present in the "normal" sera of the Australian aborigine because the aborigine never received blood transfusion, and the "antibody" is present in the sera of patients with hemophilia because they have previously received multiple blood transfusions, and tentatively called the protein as "Australia antigen" [25].

In 1968, Prince and Okochi established that the Australian antigen was found in most patients with serum hepatitis but not infectious hepatitis, and he established that blood containing Australian antigen was more likely to cause post-transfusion hepatitis [49,50].

In 1969, after further research, Blumberg et al. reported the presence of "Australian antigen" in the sera of 20% of 125 patients with acute viral hepatitis in the United States [51]. Other patients showing increased frequencies of "Australian antigen" included patients with Down syndrome who have abnormal liver function tests and evidence of hepatitis in liver biopsy, patients with post-transfusion hepatitis, and leukemic patients [51]. Based on these observations, Blumberg concluded that "the antigen is selectively associated with viral hepatitis," and "intimately associated with a virus causing hepatitis," demonstrated by its presence in infected serum [51]. He discovered that the "Isolated Australian antigen is a particle 20μm in diameter with an appearance compatible with that of a virus" [51].

Subsequent studies suggested that the Australian antigen may be the infective virus of serum hepatitis itself and reported the existence of a carrier state of the Australian antigen for nearly 20 years [52]. A technologist at Dr. Blumberg's laboratory and a patient with Down's syndrome, who were previously negative for Australian antigen, both developed hepatitis and became seropositive for Australian antigen, indicating that the latter was the etiological agent of "serum hepatitis" [1].

In 1970, with the use of immune electron microscopy, Dane et al. reported the identification of HBV with "virus-like particles about 42 nm," which "may be complete virus of Australian-Antigen-Associated Hepatitis," and "that the much more numerous 22 nm particles and long forms of Australian antigen are surplus virus-coat material" [53]. These were subsequently called "Dane particles." Subsequent studies in 1971 by Almeida et al. reported the use of detergent on Dane particles and observed there were two particles, surface and core [54]. It was revealed that the Australian antigen was a surface protein of the HBV particle (i.e. the Hepatitis B Surface Antigen [HBsAg]). The latter was shown to be noninfectious but tremendously useful as a screening tool for hepatitis B infection in blood donors and infected individuals who were likely at risk of transmitting HBV to others, such as pregnant women, health care workers, homosexuals, and drug abusers. HBsAg also served as the immunogen for Blumberg's subsequent development of hepatitis B vaccine.

After the identification of the HBV as the causative agent of serum hepatitis, subsequent hepatitis research direction was guided by several new objectives: (1) the need to develop a vaccine to protect the population against hepatitis infection, (2) the need to screen for HBsAg on blood donors to prevent post-transfusion hepatitis, and (3) the need to create a new screening test, as the method of agar gel diffusion was inadequate [48].

Hepatitis B immune globulin is used to prevent hepatitis B infection in persons without demonstrated immunity to HBV who have been exposed to the virus perinatally (i.e., infants born to HBsAg-positive mothers), by cutaneous or mucosal contact with HBsAg-positive blood or bodily fluids, or by sexual contact with a person who is positive for HBsAg (or in the case of infants younger than 12 months, by exposure to a primary caregiver in whom acute hepatitis B infection has been diagnosed) [55].

The first Hepatitis B vaccine was developed in 1969 by Blumberg and Millman, based on the understanding that the presence of antibody (anti-HBs) against the HBsAg is protective against hepatitis B infection [48]. Large amounts of HBsAg, a subunit of HBV, were extracted from sera of hepatitis B carriers [48]. Through a unique method of purifying the HBsAg and separating it from the infectious particles, Blumberg and Millman isolated HBsAg from infected blood and produced a prototype HBV vaccine for which they received a patent [48]. In 1970, Krugman also attempted to develop a hepatitis B vaccine by boiling serum containing HBsAg for 1 min and using it as a vaccine to inoculate mentally retarded children at Willowbrook. However, the boiled serum appeared to provide some but not complete protection against subsequent hepatitis exposure [48].

Based on Blumberg and Millman's concept of utilizing HBsAg purified from infected blood, Merck and its executive scientist, Maurice Hilleman, were able to manufacture sufficient quantity of hepatitis B vaccine (Heptavax B) for field trials. In 1981, the field trials conducted by Wolf-Szumness showed that this highly purified formalin-inactivated vaccine (Heptavax B) prepared from HBsAg-positive plasma was safe, immunogenic, and highly efficacious with 95% of vaccinated subjects developing antibody against HBsAg [56]. The vaccine protected against acute hepatitis B infection, asymptomatic infection, and chronic antigenemia. Within the next 2 years, FDA approved the Heptavax B.

Large-scale production of the hepatitis subunit vaccine was limited by the available supply of infected blood from hepatitis B carrier and the potential risk of transmitting other viruses present in the infected blood. In 1982, William Rutter and his colleagues were successful in synthesizing HBsAg in genetically altered yeast [57]. In 1984, Maurice Hillemen reported their success in producing this second generation of hepatitis B vaccine by recombinant monoclonal DNA techniques that allowed the expression of HBsAg in yeast cells, the first example of a vaccine produced from recombinant cells that is effective against a human infection [58]. By using recombinant yeast, the production of the hepatitis B vaccine eliminated the need for infected blood from hepatitis B carrier, and thus the potential risk of infection from HBV and other unknown infectious agents that might be present in the infected blood and the limitation of available supply of infected blood from hepatitis B carriers [58]. This recombinant vaccine was the first vaccine produced by genetic engineering for use in humans and was licensed by the FDA in 1986.

In 1991, "the Advisory Committee on Immunization Practices (ACIP) recommended a comprehensive strategy to eliminate HBV transmission in the United States, including universal vaccination of infants" using recombinant hepatitis vaccine on all newborn infants and children in the United States [55]. In 1995, the ACIP added "routine immunization of adolescents," and in 1999, "immunization of all persons up to 18 years of age" [55]. This strategy has reduced the overall annual incidence of acute HBV infections in the United States by 67%, "from 8.5 cases per 100,000 persons in 1990 to 2.8 per 100,000 in 2002" [55]. Based on data showing "a decrease in the burden of acute (and hence chronic) HBV infection as a result of immunization programs for infants, children, and adolescents," the WHO "recommended that all countries provide universal HBV immunization programs for infants and adolescents" [55]. By 2003, "79 percent of the 192 WHO member states had adopted policies of universal childhood immunization against HBV" [55]. Those countries benefited as a result of the policies; Taiwan, for example, experienced "dra-

matic decline in the incidence of neonatal HBV infections and subsequent sequelae in Taiwan after the introduction and widespread use of hepatitis B vaccine" [55]. There have also been efforts to spread the vaccine to healthcare workers: "since 1991, the Occupational Safety and Health Administration has mandated that health care workers be educated about the vaccine and that employers offer it free of charge" [55]. In 2000, it was estimated that a billion doses of the vaccine had been administered making it one of the most commonly used vaccine in the world [48].

Notably, the HBV vaccine is also a first "cancer vaccine" that can prevent hepatocellular carcinoma, the primary cancer of the liver and one of the most common cancers in the world [48]. Most of the primary liver cancer is associated with HBV or HCV infection. In addition to providing protection against HBV infection, the hepatitis B vaccine also prevents the development of the deadly hepatocellular carcinoma caused by HBV infection [48]. "The apparent success of the HBV vaccine in the prevention of primary cancer of the liver has encouraged the search of other vaccines for cancer prevention" [48].

Furthermore, over the decades since the implementation of the hepatitis B vaccine, significant drop in hepatitis B carriers in Asian countries has been observed. The current Hepatitis B vaccine demonstrates "protective serum anti-HBs antibody concentrations" in "90 percent of healthy adults and 95 percent of infants, children, and adolescents" after completion of the vaccine series [55]. For "immunocompetent persons in whom antibody levels of at least 10 mIU per millimeter develop, the efficacy of the vaccine is nearly 100 percent" [55].

POST-TRANSFUSION HEPATITIS/ TRANSFUSION-ASSOCIATED HEPATITIS

In 1943, Beeson reported seven patients who developed jaundice between 1 to 4 months after receiving whole blood or plasma transfusion [59]. Based on the fact that four of the seven patients received four or more transfusions with relatively large volumes of blood plasma, the authors suggested that the "risk of receiving a jaundice producing substance in a transfusion may be increased in proportion to the number of donors from whom blood or plasma is received" [59]. An explanation for this phenomenon is that "jaundice is caused by a virus which happened to be present in the body fluids of donors, and which, after a long incubation period, produced a hepatitis in the recipient" [59]. Unlike "campaign jaundice," these post-transfusion outbreaks of jaundice had "a long incubation period, varying from one to seven months" [59].

In the 1960s, a large percentage of donated blood originated from paid donors, who are at a higher risk to be infected with hepatitis virus than volunteer donors. Consequently, the incidence of post-transfusion hepatitis was high. In 1964, 1 year before Blumberg discovered Australian antigen, Grady et al. reported that "the apparently lower incidence of post-transfusion hepatitis in Boston can be related to the type of blood donors used," finding that "the lowest incidence of post-transfusion hepatitis was seen when commercially supplied blood was avoided" [60]. The authors recommended blood bank personnel to "exclude potential carriers of hepatitis virus by rejecting prospective blood donors with jaundice or signs or symptoms of any infectious disease" [60]. Such donors include "those having previously donated blood to persons in whom hepatitis subsequently developed, those with signs of excess exposure to potentially contaminated needles (e.g., narcotic addicts), those giving a history of infectious hepatitis and those whose appearance suggest unreliability" [60]. The authors stated that "it is not a new idea that alcoholic patients, drug addicts and other unreliable persons who deny disease for fear of rejection are attracted to any receptive and remunerative blood procurement center" [60].

The direct relationship between post-transfusion hepatitis and HBV was confirmed by Gocke et al. who, in 1970, reported, "Transfusion of blood containing Australia antigen was associated with development of hepatitis or an antibody response in 74% (31 of 42) recipients" [61]. Based on the findings,

the authors concluded that "a positive correlation does exist, and that transfusion of blood containing Australia antigen is hazardous" [61]. However, the authors cautioned that "some cases of post-transfusion hepatitis occur in recipients of antigen-negative donor blood" [61]. At the time, further work was required "to determine whether the latter observation reflects a lack of sensitivity in the test system or the existence of other infectious agents" [61]. Since then, the hepatitis B surface antigen has been used as a marker for post-transfusion hepatitis.

In 1970, Roselyn Yalow and Solomon Berson developed a serologic test called radioimmunoassay that can detect the presence of minute quantities of the HBsAg and antibody (anti-HBs) in blood [62]. This method is simpler and more sensitive than the agar gel diffusion method used by Blumberg in the discovery of Australia antigen. The ability to detect the infectious virus in the blood represented the first available method to screen the blood supply for infectious hepatitis virus.

In 1972, the FDA mandated screening of all blood donations for HBsAg. In 1976, FDA further mandated the use of an all-voluntary donor system [63]. These measures by FDA led to a substantial reduction in post-transfusion hepatitis caused by HBV [63]. HAV was determined to not be a causative agent of post-transfusion hepatitis. However, up to 10% of blood transfusion recipients continued to develop post-transfusion hepatitis, of which most cases were attributed to an unknown "non-A, non-B hepatitis" [63]. Accordingly, a "Non-A Non-B hepatitis (NANBH)" was defined as a cause of post-transfusion hepatitis after exclusion of HAV and HBV [63]. In 1986, FDA implemented the use of surrogate marker testing for NANBH and reduced the per unit risk of post-transfusion hepatitis from 1 in 200 to about 1 in 400 [63]. HCV was discovered in 1989 and was established as the causative agent of over 90% of non-A, non-B post-transfusion hepatitis [63].

DISCOVERY OF HCV

Globally, 130 to 170 million people, or about 3% of the global population, have chronic hepatitis C, a "major cause of liver cirrhosis and hepatocellular carcinoma" and "the most common cause of liver-related death and reason for liver transplantation" [64]. About 3.2 million Americans, or 1% of the US population, are infected with HCV [65]. In the United States, hepatitis C recently surpassed human immunodeficiency virus (HIV) infection as a cause of death [64].

In 1988, HCV was discovered by Michael Houghton working at a California biotechnology company, Chiron, in collaboration with the CDC [66,67]. The discovery of HCV is unique, as all other methods that led to the discovery of HAV and HBV had failed to identify the causative agent of NANBH, and novel molecular methods were used to identify the viral genome, much before visualization of the actual virus. Houghton identified a cDNA clone using sera of chimpanzees and humans with NANBH that hybridized with a single-stranded RNA made of about 10,000 nucleotides only from sera from NANBH patients. The RNA was homologous to genome of flaviviruses and was subsequently recognized as the "hepatitis C virus."

Later in 1989, Alter et al. measured the antibody (anti-HCV) to HCV, which causes non-A, non-B hepatitis, by radioimmunoassay in prospectively followed transfusion recipients and their donors [68]. Of 15 patients with chronic non-A, non-B hepatitis documented by liver biopsy, all seroconverted for the antibody; of five with acute resolving non-A, non-B hepatitis, three (60%) seroconverted [68]. The authors concluded that HCV is the predominant agent of transfusion-associated non-A, non-B hepatitis, and screening of donors for anti-HCV could prevent the majority of disease cases [68]. "Surrogate" assays for HCV infection "would have detected approximately half the anti-HCV-positive donors involved in the transmission of hepatitis that we identified" [68]. The authors also acknowledged the possibility of "a second non-A, non-B hepatitis virus" causing the absence of anti-

HCV in cases clinically diagnosed as non-A, non-B hepatitis and "estimated that the routine application of this assay in donor screening would detect approximately 85 percent of those capable of transmitting non-A, non-B hepatitis" [68]. In summary, "[measures] taken to exclude donors who are at risk for exposure to the human immunodeficiency virus, the increased use of autologous blood, and the introduction of surrogate assays have all served to diminish the risk of transfusion-transmitted hepatitis" and "[the] coming introduction of anti-HCV assay should bring a further reduction in this risk, and most important, a reduction in the long-term consequences of this common blood-borne infection" [68].

In 1990, blood screening for hepatitis C began. The missing piece to the post-transfusion hepatitis puzzle was provided by Donahue et al. in 1992 who wrote, "The most common serious complication of blood transfusion is post-transfusion hepatitis from hepatitis C virus (HCV)" [69]. This observation is supported by the sharp decrease "in the incidence of post-transfusion hepatitis C," since the "implementation of donor screening for surrogate markers (liver function tests) and antibodies to HCV" with a rate of about 3 per 10,000 units transfused [69].

The introduction of more sensitive nucleic acid amplification testing (NAT) for HCV further decreased the risk of HCV transmission through blood transfusion to approximately 1 in 2 million [70]. The risk of HCV transmission through blood transfusion is far less than that of HBV risk, "which remains at 1 in 200,000 to 500,000 using a combination of anti-hepatitis B core and hepatitis B surface antigen testing" [70].

Hepatitis C Therapy

Antiviral therapy for chronic hepatitis C first began almost 30 years ago with alpha interferon-based treatment [71]. The side effects of cytokine, the need for up to a year of therapy, and the limited response rate of 50% or less, even among carefully selected patients, limited the impact of this early therapy [71]. The recent development of direct-acting antiviral agents (DAAs) has revolutionized HCV treatment, offering prospects for the first comprehensive cure of a chronic viral infection in humans [64]. These new regimens include the combination of ledipasvir and sofosbuvir, two new direct-acting anti-viral agents with potent activity against HCV yield rates of sustained virologic response of 93% to 99% [71].

Unfortunately, challenges in combating HCV infection effectively and comprehensively remain [64]. First, patients with HCV often are diagnosed at a late stage (in high-income countries) or seldom diagnosed at all (in low- or middle-income countries) due to a lack of effective screening programs [64]. Most Americans with HCV became infected decades ago and are unaware of their status, since the symptoms have not yet manifested [72]. Half of the estimated 3.2 million Americans infected with HCV may not be aware that they are infected [71]. Rich et al. projects that in the absence of large-scale efforts at diagnosis and treatment, the burden of HCV-associated disease is expected to increase dramatically in the near future, and more than 1 million people are expected to die from HCV by 2060 in the United States [72].

Second, the high cost of DAAs hampers the effort to fight HCV [64]. A 12-week regimen of sofosbuvir alone currently costs $84,000, or $1,000 per tablet [71]. Adding ledipasvir to the treatment will further increase the cost, not to mention the expenses for diagnostic assays, monitoring, and physician visits [71]. With the current estimates of costs, treating even half of the estimated 3.2 million US residents currently infected with HCV "would add billions of dollars to an already overburdened medical care system" [71]. The high cost will limit their use in most infected patients in low- or middle-income countries and may lead to the "selective use of DAAs for certain patient subgroup" in high-income countries [64].

Third, reinfection remains a possibility even after successful curative therapy [64]. The "extraordinary sequence heterogeneity and ability to evade host immune responses" makes developing a "broadly

protective vaccine" that could potentially eradicate HCV difficult [64].

Another significant challenge stems from "poor global advocacy, perhaps due in part to a false perception of the indolent course of HCV," resulting in most patients with HCV infection in low- or middle-income counties remaining untreated [73]. Despite the fact that the "global mortality burden of viral hepatitis (A, B, C, and E) is similar to that of HIV and higher than that of tuberculosis or malaria," viral hepatitis lacks the political support, national and global policymaking and funding, and social activism in comparison with these other global diseases [73].

In the meantime, there are other ways to control HCV infection on a global scale: developing effective HCV screening programs, including full implementation of birth-cohort screening in the United States, establishing access to affordable treatment in low- and middle-income countries, and developing strategies for reducing the risk of transmission (e.g., safe injection practices) [64].

Another possibility is to target and engage "higher-prevalence countries," such as those with low-and middle-incomes, and to prioritize "higher-risk groups, such as patients with advanced liver fibrosis, and HIV or hepatitis B co-infection" for therapy [73]. Jayasekera et al. advocated "efforts toward equitable access" for the "definitive, curative therapies" to treat HCV, and urged "[listing] DAAs as essential medicine," "[creating] novel international funding streams," "[allowing] legal pathways for generic-agent manufacture," "[differential] pricing of branded originator drugs," and "[task] shifting of testing and treatment from physicians" [73].
According to Holmberg et al., "there are many points of intervention" in identification and care of patients with HCV infection which will improve the identification and care of patients with HCV and mitigate the increase in hospitalization and deaths resulting from HCV infection [65]. One example is to implement a one-time HCV screening test as recommended by the CDC for those born between 1945 and 1965, which can "help identify the many infected people who would not be targeted for testing as the result of established risk-based testing strategies" [65]. In addition, "a better job of getting HCV-infected persons who know their HCV status into care, evaluated, and, as appropriate, treated" is required [65].

Notably, an increased focus on screening and treatment for HCV in the criminal justice system is needed [72]. According to Rich et al., it is the "best place to efficiently identify and cure the greatest number of HCV-infected people," with more than 10 million Americans entering and leaving prisons and jails each year, "including nearly one of every three HCV-infected Americans" [72]. The war of drugs has led to the highest per capita incarceration rate in the world; as a result, Rich et al. reasoned, "most Americans who injected drugs have been incarcerated at some point in their lives" [72]. The rate of HCV infection in the incarcerated population has reached epidemic proportions, with one in six prisoners infected with HCV [72]. The prevalence of HCV in prisons, combined with the fact that most HCV infection in the United States is the result of past use of injection drugs and that more than 95% of prisoners are eventually released, demonstrates the effect of the incarcerated population on HCV in the community [72]. As a result, screening and treatment in the criminal justice system represents a critical opportunity to have a substantial effect on the epidemic of HCV [72]. With the low cost of screening and the high prevalence of HCV in the incarcerated population, Rich et al. advocated that everyone in that population should be screened [72]. Furthermore, early detection and treatment in correctional settings could prevent future need for treatment, which, along with its attendant costs, would occur predominantly in the community while also preventing the spread of HCV [72].

Discovery of Hepatitis D Virus

In 1977, Rizzetto et al. described a new antigen, termed δ (delta), detected by direct immunofluorescence in the liver cell nuclei and in the blood of patients with HBsAg-positive chronic liver disease

[74]. They also reported that δ antibody was found only in the serum of chronic HBsAg-positive carriers, with a high prevalence in patients with liver damage [74]. Subsequently, it was determined that delta antigen was not part of HBV but of a separate defective virus that requires the presence of HBV for infection and was named the hepatitis D virus (HDV) to conform to the nomenclature of hepatitis viruses and classified within the genus of Deltavirus [75].

The superinfection of the delta virus in patients with hepatitis B resulted in an increased level of inflammation and necrosis of liver cells, and hence a more rapidly progressive form of HBV-related chronic liver disease, including liver cirrhosis, liver decompensation, and death [76]. Currently, HDV infection is distributed worldwide, involving approximately 5% of HBsAg carriers amounting to 15 to 20 million HDV-infected individuals [76]. It is highly endemic in Mediterranean countries, northern parts of South America, and Central Africa [76]. The virus is transmitted through the parental route and is associated with intravenous drug use, multiple sexual partners, tattooing, and piercing [76]. However, currently there is no efficient therapy except for prolonged treatment with recombinant interferons, which is the only therapy that has shown antiviral activity against HDV, but with only 20–40% efficiency [76].

DISCOVERY OF HEV

In 1980, a "common source waterborne epidemic of viral hepatitis" associated with gross fecal contamination of water was reported in Kashmir valley [77]. The epidemic's waterborne nature established the "nonparenteral mode of spread, i.e. a fecal oral spread of the disease" [77]. However, serologic tests for hepatitis A and hepatitis B failed to "reveal an etiologic agent of hepatitis," even though "the mode of spread of the epidemic, length of incubation, clinical features and biochemical test results of the patients studied resembled that of hepatitis A" [77]. Therefore, the results suggested the possibility of another human hepatitis virus distinct from hepatitis A or hepatitis B.

Another report in 1980 found similar results and described "one of the first serologically documented reports of epidemic hepatitis transmitted via contaminated water that was not caused by HAV" [78]. The author stated that "a large portion of hepatitis in India seems to be caused by previously unrecognized agents," signifying "growing evidence that non-A, non-B agents which epidemiologically resemble HAV exist," and are "responsible for much morbidity and mortality, especially in parts of Asia" [78].

In a study by Balayan, a human volunteer was orally fed "pooled stool extracts from patients with presumed non-A, non-B hepatitis" [79]. The volunteer consequently developed non-A, non-B hepatitis, with symptoms similar to those of hepatitis A, without "serological evidence of recent hepatitis A or hepatitis B infection" [79]. Examining the volunteer's stool using immune electron microscopy revealed "27- to 30-nm spherical virus-like particles indicating possible causative agent of the fecal-oral non-A, non-B hepatitis" [79]. In 1993, Chauhan et al. reported the presence of HEV in both the stools and sera of one of the authors who deliberately infected himself by orally ingesting stools from a patient infected from the hepatitis E epidemic, hence suggesting the possibility of sporadic transmission of HEV parenterally [80].

Currently, hepatitis E, the fifth known form of human viral hepatitis, is probably "the most common cause of acute hepatitis and jaundice in the world" [81]. It is the second most common form of human viral hepatitis in the United States [81]. It is less common in the United States and other developed nations than in developing countries in Asia and Africa, where it is a "major public health problem" [82]. According to population-based surveys from 1988 to 1994, 21.0% of US adults had anti-HEV antibody, which is "lower than that with anti-HAV antibody (38.3%), but higher than that with antibodies against hepatitis B (5.7%) or hepatitis C (2.0%)" [81]. An estimated one-third of the world's population has been infected with HEV, based on seroprevalence [82]. The lifetime infection risk is

more than 60% in India, with hundreds of thousands of illnesses every year [82].

The role of hepatitis E in causing liver disease is currently not well known. Hepatitis E is usually self-limited and can occur sporadically and in epidemics [82]. However, for pregnant women, who have "the highest risk of associated hepatic failure," the case fatality ratio increases from 5% to 25%, and those who survive often have "high rates of spontaneous abortion and stillbirth" [82]. Interestingly, similar epidemiological and clinical presentation was reported in 1863 by Harley on the island of Martinique, where an epidemic "inflicted thirty pregnant women" and resulted in twenty fatalities after "an abortion or premature labor" [9]. Currently, the rHev vaccine is effective in preventing hepatitis E, with a reported efficacy of 95.5% [82].

DISCOVERY OF HEPATITIS G VIRUS

Approximately "10 to 15 percent of patients with parenterally transmitted or transfusion-associated non-A, non-B hepatitis have no evidence of hepatitis C virus (HCV) infection" and were "classified as having non-A-E hepatitis" [85,86]. In 1996, a new member of the Flaviviridae family, the hepatitis G virus (HGV), was discovered. This virus was identical to another newly cloned agent designated as GB virus, type C [87,88]. However, additional studies did not substantiate HGV as an etiologic agent of non-A-E hepatitis [85]. Persistent infection with HGV was common, but it did not lead to chronic disease and did not affect the clinical course in patients with hepatitis A, B, or C [85]. Other studies of patients with transfusion-associated hepatitis concluded that "HGV was common in a group of volunteer blood donors, and it [could] be transmitted by transfusion" but "most HGV infections were not associated with hepatitis" and "did not worsen the course of concurrent HCV infection" [86]. The authors found "no causal relation between HGV and hepatitis" [86].

ONGOING CHALLENGES OF A GLOBAL EPIDEMIC

The advances in our knowledge of hepatitis are breathtaking and miraculous. Through the work of scientists such as Dr. Baruch Blumberg, we have greatly advanced our knowledge of viral hepatitis. The development of vaccines and cures or other therapies for certain types of viral hepatitis provides hope for the future. As of 2009, 91% of WHO Member States included the HBV vaccine in their infant immunization programs and more than 70% of infants received three doses of this vaccine, which provided them with livelong protection from HBV [89]. 179 Countries have introduced the HBV vaccine that has prevented approximately 1.3 million deaths worldwide [89].

Despite this progress, hepatitis is still a significant global epidemic infecting 1 in 12 people worldwide with hepatitis B infecting about 2 billion and hepatitis C infecting about 150 million people [89]. It is estimated that about 1.4 million new hepatitis A virus infections and 20 million hepatitis E infections occur globally each year [89]. In many countries, there is a dearth of adequate knowledge and awareness among both the general population and health professionals concerning hepatitis, according to the WHO [89]. Countries without a hepatitis surveillance system have difficulty making evidence-based policy decisions [89]. Another barrier to comprehensive treatment of hepatitis is the failure to screen donated blood for diseases such as HBV, HCV, HIV, and syphilis [89]. Irregular supply of test kits and the high costs of HCV screening tests are among the most commonly reported barriers to screening of donated blood [89]. Furthermore, the lack of access to both clean drinking water and improved sanitation, especially in low-income countries, makes it easy for large portions of a population to contract hepatitis [89]. Although treatment for viral hepatitis B and C exists, it is often inaccessible or too expensive for most people in settings where resources are scarce [89].

In order to fight hepatitis, the WHO advocates raising awareness of viral hepatitis infections in order to educate at-risk populations. Other recommended

measures include hepatitis policy based on accurate data, and prevention of transmission, through immunization policies and protecting high-risk groups [89]. Guidelines for screening, for increasing access to care, for treating patients with chronic HBV or HCV infection, and for managing drug resistance are also critical to fighting hepatitis [89].

CONCLUSIONS

This article summarized the historical milestones in the discovery of viral hepatitis. The combination of knowledge of the history of viral hepatitis with effective policies that raise global awareness; prevent transmission; and provide accessible and affordable screening, care, and treatments to those infected with viral hepatitis will allow us to begin effectively combating this silent global epidemic.

Abbreviations: HAV, hepatitis A virus; HBV, hepatitis B virus; HCV, hepatitis C virus; NANBH, Non-A, Non-B hepatitis; HDV, hepatitis D virus, HEV, hepatitis E virus, HFV, hepatitis F virus; HGV, hepatitis G virus; HBsAg, hepatitis surface antigen; Anti-HBs antibody, Anti-hepatitis B surface antigen antibody; WHO, World Health Organization; WSS, Willowbrook State School; ACIP, Advisory Committee on Immunization Practices; FDA, US Food and Drug Administration; HIV, human immunodeficiency virus.

REFERENCES

1. Zuckerman AJ (1983) The History of Viral Hepatitis from Antiquity to the Present. In Deinhardt F, Deinhardt J, editors, *Viral Hepatitis: Laboratory and Clinical Science*, New York, Marcel Dekker, 3-32.
2. Martin NA (2003) The Discovery of Viral Hepatitis: A Military Perspective. *J R Army Med Corps* 149: 121-4.
3. Starr D (1998) *Blood. An Epic History of Medicine and Commerce*. Perennial. Harper Collins Publishers.
4. (1940) Zacharis. In: *The Letters of Saint Boniface (751AD)*. Translated by E. Emetson. New York: Columbia University Press, 159-64.
5. Krugman S. Viral Heptatis (1976) Overview and Historical Perspectives. *Yale J Biol Med* 49:199-203.
6. Monro D (1764) An Account of Diseases Which Were Most Frequent in British Military Hospitals in Germany from January 1761 to the Return of the Troops in England in March 1763, Millar and Cadell, London, 206-15.
7. Girddlestone T (1787) Essays on the Hepatitis and Spasmodic Affections in India; Founded on Observations Made Whilst On Service With His Majesty's Troops in Different Parts of That County. S. Murray, London, 1-25.
8. McClean C (1818) Practical Illustrations of the Progress of Medical Improvement, For the Last Thirty Years; or Histories of Acute Diseases as Fever, Dysentry, Hepatitis, and Plaque. R & R Gilbert, London, 48-52.
9. Harley G (1863) Jaundice: Its Pathology and Treatment. With the Application of Physiological Chemistry to the Detection and Treatment of Disease of the Liver and Pancreas. Walton, London, 90-100.
10. Harley G (1883) The Diseases of the Liver: With or Without Jaundice, With the Special Application of Physiological Chemistry to the Diagnosis and Treatment. J & A Churchill, London, 333-63.
11. Virchow R (1865) Uber das Vorkommen und den Nachweiss des Hepatogenen, insbesondere des Katarrhalischen Icterus. *Virchows Arch Path Anat Physiol* 32: 117-25.
12. Murchison C (1885) Clinical Lectures on Diseases of the Liver, Jaundice and Abdominal Dropsy, 3rd ed., William Wood and Co., New York.
13. (1888) Medical and Surgical History of the War of the Rebellion. Medical History. Washington: Government Printing Office, 1(3): 874-79.
14. Martin CJ (1918) Concerning the Pathology and Aetiology of the Infectious Jaundice Common at the Dardanelles, 1915. *J R Army Med Corps* 30: 102-9.
15. Macpherson WG, Harringham WP, Elliot TR, Balfour A (ed.) (1921) Epidemic Catarrhal Jaundice, in History of the Great War Based on Official Documents. Medical Services. *Diseases of the War*, HMSO, London, 1:395-400.
16. Havens WP (1961) Viral Hepatitis. *Yale J Biol Med* 34: 314-28.
17. Gutzeit K (1950) Die Hepatitis Epidemica, Munchen Med. Wochenschr 92: 1161, 1295.
18. Cullinan ER (1952) Chapter IX. Infective Hepatitis. In Cope YZ, editor, Medical History of the Second World War. Medicine and Pathology; HMSO, London, 230-49.
19. MacCallum FO, Bradley WH, Van Rooyen CE, Sheehan HL (1952) Chapter IX. Infective Hepatitis. In Cope YZ, editor, Medical History of the Second World War. Medicine and Pathology; HMSO, London, 252-67.
20. Stokes J, Jr., Neefe FR (1945) The Prevention and Attenuation of Infectious Hepatitis by γ-Gobulin. *JAMA* 127: 144-5.
21. Craig AS, Shaffner, W (2004) Prevention of Hepatitis A with the Hepatitis A Vaccine. *N Eng J Med* 350: 476-81.
22. Feinstone SM, Kapikian AZ, Purcell RH (1973) Hepatitis A. Detection by Immune Electron Microscopy of a Virus-Like Antigen Associated with Acute Illness. *Science* 182:1026-8.
23. Victor JC, Monto AS, Surdina TY, Suleimenova SZ, Vaughan G, Nainan OV, et al. (2007) Hepatitis A Vaccine versus Immune Globulin for Postexposure Prophylaxis. *New Eng J Med* 357: 1685-94.
24. Bar-Gal GK, Kim MJ, Klein A, Shin DH, Oh CS, Kim JW, et al. (2012) Tracing Hepatitis B Virus to the 16th Century in a Korean Mummy. *Hepatology* 56:1671-80.
25. Blumberg B, Alter HJ, Visnich S (1965) A "New" Antigen in Leukemic Sera. *JAMA* 191: 541-6.
26. Lürman A (1885) Eine Icterus Epidemic. Berlin Klin. Wochenschr, 22:207.
27. McDonald S (1908) Acute Yellow Atrophy. *Edinburgh Med J* 1(NS):83-8.

28. Stokes JH, Ruedemann R, Jr., Lemon WS (1920) Epidemic Infectious Jaundice and Its Relation to the Therapy of Syphilis. *Arch Int Med* 26:521-43.
29. MacCallum FO (1943) Jaundice in Syphilitics. *Br J Vener Dis* 19:63.
30. Findlay GM, MacCallum FO (1937) Note on Acute Hepatitis and Yellow Fever Immunization. *Tran Roy Soc Trop Med & Hyg* 31:297-308.
31. Propert SA (1938) Hepatitis after prophylactic serum. *BMJ* 2:677-8.
32. Editorial (1942) Jaundice Following Yellow Fever Vaccination. JAMA 119:1110.
33. Turner RH, Snavely JR, Grossman EB, Buchanan RN, Foster SO (1944) Some Clinical Studies of Acute Hepatitis Occurring in Soldiers After Inoculation With Yellow Fever Vaccine, With Especial Consideration of Severe Attacks. *Ann Int Med* 20:193-218.
34. Seefe LB, Beebe GW, Hoofnagle JH, Norman JE, Buskell-Bales Z, Waggoner JG, et al. (1987) A serologic follow-up of the 1942 epidemic of post-vaccination hepatitis in the United States Army. *N Eng J Med* 316: 965-70.
35. MacCallum FO (1947) Homologous serum jaundice. *Lancet* 2:691-2.
36. Voegt H (1942) Zur Aetiologie der Hepatitis Epidemica. *München med Wchnschr* 89:76.
37. Cameron JDS (1943) Infectious Hepatitis. *Quat J Med* 12:139-55.
38. MacCallum FO, Bradley WH (1944) Transmission of Infectious Hepatitis to Human Volunteers. *Lancet* 2:228.
39. Paul JR, Havens, WP (1945) Transmission Experiments in Serum Jaundice and Infectious Hepatitis. *JAMA* 128:911-5.
40. Neefe JR, Stokes J Jr, Reinhold JG (1945) Oral Administration to Volunteers of Feces from Patients with Homologous Serum Hepatitis and Infectious (Epidemic) Hepatitis. *Am J Med Sci* 210:29-32.
41. Neefe JR, Stokes J Jr. (1945) An Epidemic of Infectious Hepatitis Apparently Due to a Water Borne Agent. *JAMA* 128:1063-75.
42. Havens WP, Jr. (1946) Immunity in Experimentally Induced Infectious Hepatitis. *J Exp Med* 84(5): 403-6.
43. Neefe JR, Gellis SS, Stokes J. (1946) Homologous Serum Hepatitis and Infectious (Epidemic) Hepatitis: Studies in Volunteers Bearing on Immunological and Other Characteristics of the Etiological Agents. *Am J Med* 1:3-22.
44. Stokes J, Farquhar JA, Drake ME, *et al*. (1951) Length of Protection by Immune Serum Globulin (γ-Globulin) During Epidemics. *JAMA* 147:714-9.
45. Beecher HK (1966) Ethics and Clinical Research. *New Eng J Med* 274:1354-60.
46. Rothman DJ (1982) Were Tuskegee & Willowbrook "Studies in Nature"? *The Hastings Center Report* 12(2): 5-7.
47. Krugman S (1986) The Willowbrook Hepatitis Studies Revisited: Ethical Aspects. *Rev Infect Dis* 8(1):156-62.
48. Blumberg BS (2002) *Hepatitis B: The Hunt for a Killer Virus*. Princeton and Oxford, Princeton University Press, 72-146.
49. Okochi K, Murakami S (1968) Observations on Australian antigen in Japanese. *Vox Sang* 15:374-85.
50. Prince AM (1968) An Antigen Defected in the Blood During the Incubation Period of Serum Hepatitis. *Proc Natl Acad Sci USA* 60:814-21.
51. London WT, Sutnick AI, Blumberg B (1969) Australia Antigen and Acute Viral Hepatitis. *Ann Int Med* 70:55-9.
52. Zuckerman AJ, Taylor PE (1969) Persistence of the Serum Hepatitis (SH-Australia) Antigen for Many Years. *Nature* 223:81-2.
53. Dane DS, Cameron CH, Briggs M (1970) Virus-like Particles in Serum of Patients with Australia-Antigen-Associated Hepatitis. Lancet 1:695-8.
54. Almeida JD, Rubenstein D, Stott EJ (1971) New Antigen-Antibody System in Australia-Antigen-positive Hepatitis. *Lancet* 2:1225-6.
55. Poland GA, Jacobson RM (2004) Prevention of Hepatitis B with the Hepatitis B Vaccine. *New Eng J Med* 351:2832-8.
56. Szmuness W, Stevens CE, Zang EA, Harley EJ, Kellner A (1981) A Controlled Clinical Trial of the Efficacy of the Hepatitis B Vaccine (Heptavax B): A Final Report. *Hepatology* 1(5): 377-84.
57. Valenzeula P, Medina A, Rutter WJ, Ammerer G, Hall BD (1982) Synthesis and Assembly of Hepatitis B Virus Surface Antigen Particles in Yeast. *Nature* 298: 347-50.
58. McAleer WJ, Buynak EB, Maigetter RZ, Wampler DE, Miller WJ, Hilleman MR (1984) Human Hepatitis B Vaccine from Recombinant Yeast. *Nature* 307:178-80.
59. Beeson PB (1943) Jaundice Occurring One to Four Months After Transfusion of Blood or Plasma. Report of Seven Cases. *JAMA* 121:1332-4.
60. Grady GF, Chalmers TC (1964) Risk of Post-Transfusion Viral Hepatitis. *N Eng J Med* 271: 331-42.
61. Gocke DJ, Greenberg HB, Kavey NB (1970) Correlation of Australia Antigen with Post-Transfusion Hepatitis. JAMA 212:877-9.
62. Walsh JH, Yalow RS, Berson SA (1970) Radioimmunoassay of Australia Antigen. *Vox Sang* 19(3):217-24.
63. Tobler LH, Busch MP (1997) History of Post-Transfusion Hepatitis. *Clinical Chemistry* 43:1487-93.
64. Chung RT, Baumert TF (2014) Perspective: Curing Chronic Hepatitis C – The Arc of a Medical Triumph. *New Eng J Med* 370(17):1576-8.
65. Holmberg SD, Spradling PR, Moorman AC, Denniston MM (2013) Hepatitis C in the United States. *New Eng J Med* 368(20):1859-61.
66. Ezzell C (1977) Candidate Cause Identified of Non-A, Non-B Hepatitis [News]. *Nature* 333:195.
67. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M (1989) Isolation of a cDNA Clone Derived from a Blood-Borne Non-A, Non-B Viral Hepatitis Genome. *Science* 244:359-61.
68. Alter HJ, Purcell RH, Shih JW, Melpolder JC, Houghton M, Choo QL, et al. (1989) Detection of Antibody to Hepatitis C Virus in Prospectively Followed Transfusion Recipients With Acute and Chronic Non-A, Non-B Hepatitis. *New Eng J Med* 321: 1494-500.
69. Donahue JG, Munoz A, Ness PM, Brown DE, Yawn DH, McAllister HA, et al. (1992) The Declining Risk of Post-

Transfusion Hepatitis C Virus Infection. *N Eng J Med* 327:369-73.
70. Stramer SL (2007) Current Risk of Transfusion-Transmitted Agent. *Arch Path Lab Med* 131: 702-7.
71. Hoofnagle JH, Sherker AH (2014) Editorial: Therapy for Hepatitis C – The Costs of Success. *New Eng J Med* 370 (16):1552-3.
72. Rich JD, Allen SA, Williams BA (2014) Responding to Hep C through the Criminal Justice System. *New Eng J Med* 370(20):1871-4.
73. Jayaskera CR, Barry B, Roberts LR, Nguyen MH (2014) Perspective: Treating Hepatitis C in Lower-Income Countries. *New Eng J Med* 370(20):1869-71.
74. Rizzetto M, Canese MG, Arico S, Crivelli O, Trepo C, Bonino F, et al. (1977) Immunofluorescence Detection of a New Antigen/Antibody System (Delta/Anti-Delta) Associated with Hepatitis B Virus in Liver and Serum HBsAg Carriers. *Gut* 18:997-1003.
75. Rizzetto M (2009) Hepatitis D: Thirty Years After. *J Hepatol* 50:1043-50.
76. Alves C, Branco C, Cunha C (2013) Hepatitis Delta Virus: A Peculiar Virus. *Advances in Virol* 2013:560105.
77. Khuroo MS (1980) Study of an Epidemic of Non-A, Non-B Hepatitis. Possibility of Another Human Hepatitis Virus Distinct from Post-Transfusion Non-A, Non-B Type. *Am J Med* 68:818-24.
78. Wong D, Purcell RH, Sreenivasan MA (1980) Epidemic and Endemic Hepatitis in India: Evidence For a Non-A, Non-B Hepatitis Virus Aetiology. *Lancet* 316:876-9.
79. Balayan MS, Andjaparidze AG, Savinskaya SS, Ketiladze ES, Braginsky DM, Savinov AP, et al. (1983) Evidence for a Virus in a Non-A, Non-B Hepatitis Transmitted via the Fecal-Oral Route. *Intervirology* 20:23-31.
80. Chauhan A, Jameel S, Dilawari JB, Chawla YK, Kaur U, Ganguly NK (1993) Hepatitis E Virus Transmission to a Volunteer. *Lancet* 341(8838):149-50.
81. Hoofnagle JH, Nelson KE, Purcell RH (2012) Current Concepts: Hepatitis E. *New Eng J Med* 367(13):1237-44.
82. Shrestha MP, Scott, RM, Joshi DM, Mammen MP, Thapa GB, Thapa N, et al. (2007) Safety and Efficacy of a Recombinant Hepatitis E Vaccine. *New Eng J Med* 356:895-903.
83. Deka N, Sharma MD, Mukerjee R (1994) Isolation of The Novel Agent From Human Stool Samples that is associated with Sporadic non-A, non-B Hepatitis. *J Virol* 68:7810-15.
84. Bowden S (2001) New Hepatitis Viruses: Contenders and Pretenders. *J Gastroenterol Hepatol* 16(2):124-31.
85. Alter MJ, Gallagher M, Morris TT, Moyer LA, Meeks EL, Krawczynski K, et al. (1997) Acute non-A-E Hepatitis in the United States and The Role of Hepatitis G Virus Infection. *N Eng J Med* 336:741-6.
86. Alter HJ, Nakatsuji Y, Melpolder J, Wages J, Wesley R, Shih W-K, et al. (1997) The Incidence of Transfusion-Associated Hepatitis G Virus Infection and Its Relation to Liver Disease. *N Eng J Med* 336:747-54.
87. Linnen J, Wages J, Zhang-Keck ZY, Fry KE, et al. (1996) Molecular Cloning and Disease Association of Hepatitis G Virus: A Transfusion-Transmissible Agent. *Science* 271:505-8.
88. Masuko K, Mitsui T, Iwano K, Yamazaki C, Okuda K, Meguro T, et al. (1996) Infection with Hepatitis GB Virus C in Patients on Maintenance Hemodialysis. N*ew Eng J Med* 334:1485-90.
89. (2012) World Health Organization. Prevention & Control of Viral Hepatitis Infection: Framework for Global Action. Retrieved from http://www.who.int/hiv/pub/hepatitis/Framework/en/.

HMSR

VIEWPOINT
Reflection

Alma Mater Revisited: Teaching Medicine as a Fulbright Scholar

Gyorgy Baffy, MD, PhD[1]

[1]Department of Medicine, VA Boston Healthcare System, Harvard Medical School, Boston, MA 02115, USA
Correspondence should be addressed to G.B. (gbaffy@partners.org)

The prestigious Fulbright Program sends US grantees to teach or conduct research in over 155 countries. A Fulbright scholar from Harvard summarizes his recent teaching experience at a medical school in Hungary, where he graduated over 30 years ago. The one-semester fellowship provided meaningful lessons in developing and delivering a course on obesity to an internationally diverse group of students.

My first semester in the new academic year promised to be quite out of the ordinary. I packed my bags and planned to spend 3 months in Hungary teaching at the medical school where I graduated in 1980. This was the culmination of a year-long process during which I successfully applied for a Fulbright US scholarship and proposed an elective course entitled "Comprehensive Review of Obesity and Associated Disorders," to be offered within the English language program of graduate medical education at the University of Debrecen.

The Fulbright Program was founded in 1946 and has since become one of the largest international exchange programs in the world. The program is overseen by the J. William Fulbright Foreign Scholarship Board under the sponsorship of the Department of State's Bureau of Educational and Cultural Affairs with the support of national Fulbright commissions or foundations in most participating countries. There are more than 325,000 "Fulbrighters" who either came to the United States or, as in my case, went abroad to study, teach, and conduct research in all fields of arts, sciences, and public affairs. The Fulbright Program has had a tremendous impact on international collaboration, and many Fulbright alumni have occupied key roles in academia, industry, and government [1].

The Core US Scholar Award is one of the numerous bilateral teaching and research opportunities provided by the Fulbright Program. Between 2006 and 2014, a total of 6,631 such awards were granted, and 26 of these went to Harvard faculty members; in the current academic year alone, Harvard received ten additional scholarships. Within the 8-year period, there were 186 scholarships funded in Medical Sciences and 106 in Public and Global Health, of which 10 projects were related to various aspects of obesity and weight management. During this time, Hungary received 97 Fulbright scholars, including seven in the healthcare profession.

When I learned that the Hungarian-American Fulbright Commission would not only condone but in fact also encourage the idea of teaching a semester at my home university, I decided to apply. There was also little doubt as to the topic of the course: obesity was a clear choice based on my long-time clinical and research interest in fatty liver disease, mitochondrial bioenergetics, metabolic reprogramming of cancer cells, and other pathophysiological aspects of sustained nutrient excess. Obesity's pres-

ence in our current society is staggering. Furthermore, its complex nature relates to most facets of medicine. This choice enabled me to take advantage of my own clinical skills and research expertise while developing an interdisciplinary teaching program.

The application guidelines asked for a syllabus proposal specifying the course agenda, educational objectives, timetable, and evaluation methods, along with two sample seminars with major teaching points and recommended literature. I submitted my application to the Fulbright Program in the summer of 2013 for a starting date in September 2014. My objective was to focus on current advances and generate as much interest in the classroom as possible. Most of the topics were strongly related to my own clinical and research expertise and have been presented at various courses many times. For the sake of comprehensiveness, I included lesser-visited titles such as neurological control of appetite regulation and metabolic surgery (**Table 1**). The course was deliberately designed to be different from the traditional electives, with the overarching goal of reviewing every possible aspect of obesity, from molecular biology all the way to societal implications. Obesity and its adverse health effects have become a pervasive problem in most parts of the world. Why is this trend growing? What are the long-term consequences of obesity? What can we do to prevent further escalation of the obesity epidemic? Health care professionals around the world are forced to confront these questions. Through this project, I aimed to help students understand these challenges, acquire the skills to prevent and manage adverse outcomes of obesity, and disseminate these values upon returning to their countries of origin.

The medical school at the University of Debrecen is one of four similar institutions in Hungary. There are almost 1,000 students at any given time in Debrecen who attend the English language curriculum based on the traditional 6-year structure of medical education. This has essentially doubled the number of medical students in Debrecen compared to the 1970s when I attended the school. The English program has become particularly popular after Hungary joined the European Union, in which university diplomas enjoy mutual recognition in member countries. Class sizes have grown accordingly, and the current faculty/student ratio is now less than originally envisioned. The European Union has also made a great impact on the campus by funding an unprecedented pace of development in areas such as clinical genomics and stem cell therapy.

Table 1. Agenda for Course: Comprehensive Review of Obesity and Associated Disorders

	Obesity Course Agenda
1	Epidemiology and global health impact of obesity
2	Obesity, the metabolic syndrome, and cardiovascular disease
3	Principles of energy homeostasis
4	Molecular pathways of appetite regulation
5	Mechanisms of adipogenesis and ectopic fat accumulation
6	Insulin resistance and type 2 diabetes
7	Gut microbiota and obesity
8	Nonalcoholic fatty liver disease
9	Obesity and cancer
10	Obesity, complex systems science, and network medicine
11	Societal initiatives on obesity prevention and weight management
12	Nutritional counseling and lifestyle modification in obesity
13	Pharmacotherapy for obesity
14	Principles and practice of advanced obesity surgery
15	Wrap-up and perspectives

The Department of Medicine was quite supportive of the project, creating a webpage for the course where all documents, including the syllabus, timetable, student choice papers, and 20-25 full-text articles for each topic were accessible to the students with password protection. The website was updated after each class with my handouts and the attend-

ance records. The start was a bit slow, as many students signed up at the last minute. I did not receive official leave for my Fulbright fellowship and had to avoid contiguous absence from my workplace for extensive periods of time. Accordingly, I developed a rather complicated timetable in which we had two classes in some weeks and none in others, depending on whether I was in town or not. As a result, I traveled back and forth between Boston and Hungary four times during the semester. It was therefore a limited imitation of a sabbatical, but sticking with the convoluted classroom schedule still worked out well even with interruptions of up to 2 weeks.

In the first 2 weeks, a total of 32 students from 17 countries (Bangladesh, Belarus, Brazil, Ghana, Kenya, Malaysia, Nigeria, Norway, Iceland, Iran, Israel, Italy, Saudi Arabia, Sweden, Uganda, Ukraine, and the United States) signed up for my elective course. Most were in the fourth or fifth year of their medical curriculum. Teaching such a multiethnic group was a challenging but terrific experience. Cultural differences between representatives of various geographical areas were palpable, as Scandinavian students were generally outspoken and quite active, while those from West Africa remained more reserved and participated in the classroom discussions only if encouraged to do so.

During classes, I followed a teaching model based primarily on our GI pathophysiology seminars in the HST.120 program at Harvard. In the first hour, I gave a PowerPoint lecture on each topic followed by the students presenting their mini-seminars, 15 minutes each, about a paper selected from a list of original contributions in leading journals of the

"Tempus Fugit" by Erin Dvorak
Art Coordinator / Sanborn Western Camps
erindvor@gmail.com

field. I ended each talk with a few "concepts recap" slides, just to reinforce the most important definitions and to give the students another opportunity for interaction with the material or further clarification if necessary. I also encouraged the students to come up with a few pieces of "obesity wisdom," which often proved to be the focal point of our classroom discussions. It was generally understood that no matter where they were from, students would eventually face the challenges of obesity when they returned to their homeland or practice elsewhere. Thus, parallel curves of growing obesity prevalence in the US across various geographic locations, ethnicity, or socioeconomic status indicate globally similar trends and may well apply to other countries and communities [2]. All agreed that it might be just a matter of time before these curves grow sufficiently high to put obesity on the radar screen even in the leanest of countries.

Although students were initially reluctant to adopt the journal club format, ultimately, their presentations became increasingly comprehensive and elaborate. Some of them closed their talks with a cartoon relevant to the topic, a trick that I often used to boost the attention of my audience. Watching and listening to the students during these mini-seminars was perhaps the most revealing of their personal qualities and cultural diversity. The PowerPoint talks represented a wide range of skills and creativity, often defying my expectations based on prior classroom activity. Snappy, funny, and visually rich student presentations originated from every corner of the world.

There were ample opportunities for real-time teaching. Fortuitously during our course, the FDA approved a new obesity drug (naltrexone/ bupropion), and the glucagon-like peptide 1 receptor agonist liraglutide was slated for approval to treat obesity shortly thereafter [3,4]. In addition, the city of Berkeley passed the first soda tax in the US [5]. We monitored the weekly blog of Dr. David Allison from the University of Alabama compiling recently published articles and news pieces about obesity [6], watched a video on weight bias of health care workers produced by the Rudd Center for Food Policy and Obesity of Yale University [7], and watched a YouTube edition of the plenary talk by Drs. Steven Woloshin and Lisa Schwartz from Dartmouth College on the risk of over-diagnosis in medicine [8]. We spent time discussing the risks and benefits of having an enormous amount of information about obesity on the Internet and the challenges of filtering this information for our patients and ourselves.

In the final class, we spent an hour on informal student feedback, discussing what the students liked and disliked most about the course. This was probably the liveliest conversation of the entire semester. To my great relief, no one criticized or complained about my hectic timetable. The students primarily appreciated the multidisciplinary nature of the course in which we combined bench research findings with advances in medicine, surgery, and health care policy. This stood in contrast to their other courses, whose structure was more rigid. They also enjoyed discussing the latest advances rather than having a textbook-style recapitulation of what is known about obesity. The best reward was that the students asked if I was planning to teach another elective in the next academic year, and if so, they promised to sign up without hesitation.

Regrettably, even if I could qualify for a Fulbright award the second time, my vacation leave balance would not accommodate such an ambitious plan in the foreseeable future. Therefore, I will have to return to the earlier routine of giving lectures and seminars ad hoc whenever I have a chance to visit the medical school in Debrecen. The course has taught me many things of which I will also benefit when teaching at HMS. Finally, I understand now much better why there are so few of us in the medical profession who may benefit from the rewarding experience of a Fulbright scholarship, which requires a time commitment difficult to meet. Nevertheless, it is a challenge worth trying.

The Fulbright Program can be accessed via http://www.cies.org, which provides an overview of the grant opportunities and accepts applications.

REFERENCES

1. (2014) US Department of State Bureau of Educational and Cultural Affairs. The Fulbright program. Retrieved from http://eca.state.gov/fulbright/about-fulbright/j-william-fulbright-foreign-scholarship-board-ffsb.
2. Sturm R, An R (2014) Obesity and economic environments. *CA Cancer J Clin* 64:337-50.
3. US Food and Drug Administration (2014) FDA approves weight-management drug Contrave. Retrieved from http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm413896.htm.
4. US Food and Drug Administration (2014) FDA approves weight-management drug Saxenda. Retrieved from http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm427913.htm.
5. Isidore C (2014) CNN Money. New Berkeley soda tax costs 68 cents per two liter bottle. Retrieved from http://money.cnn.com/2014/11/05/news/economy/berkeley-soda-tax/.
6. Allison DB (2015) Obesity and energetics offerings. University of Alabama. Retrieved from http://www.obesityandenergetics.org/.
7. (2013) The Rudd Center for Food Policy & Obesity at Yale University. Videos exposing weight bias. Retrieved from http://www.yaleruddcenter.org/what_we_do.aspx?id=254ity.
8. Woloshin S, Schwartz L (2013) Preventing Overdiagnosis: Winding back the harms of too much medicine. Dartmouth College. Retrieved from http://www.youtube.com/watch?v=nvhs10XnsoI.

HMSR

VIEWPOINT
Health Policy

Increasing Drug Costs in the United States: Time for Reform?

Janani Panchalingam, MPH[1]; Madhu Mazumdar, PhD[1]; and Jashvant Poeran, MD, PhD[1]

[1]Institute of Healthcare Delivery Science, Department of Health Evidence & Policy, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

Correspondence should be addressed to J. Poeran (jashvant.poeran@mountsinai.org)

The ever-increasing burden of US healthcare costs—currently 18% of GDP—remains a priority for policymakers, as illustrated by the numerous efforts that have been initiated to reduce costs. However, relatively little attention has been paid to the relationship between drug approval processes or pricing strategies and the cost of drugs in the US, currently the highest in the world. Given the significant potential for cost savings in this field, federal legislation could play an important role in pharmaceutical regulation. Here, we will discuss cost-effectiveness criteria and their potential for introduction into the US drug approval process.

INTRODUCTION

The ever-increasing burden of US healthcare costs—currently 18% of GDP—remains a priority for policymakers, as illustrated by the numerous efforts that have been initiated to reduce costs. However, relatively little attention has been paid to the relationship between drug approval processes or pricing strategies and the cost of drugs in the US, currently the highest in the world. Given the significant potential for cost savings in this field, federal legislation could play an important role in pharmaceutical

regulation. Here, we will discuss cost-effectiveness criteria and their potential for introduction into the US drug approval process.

MECHANISMS OF COST DIFFERENCES

Pharmaceutical spending per capita in the US has increased from approximately $540 in 2000 to $1,010 in 2012, compared to $300 in 2000 to $498 in 2012 for the Organization for Economic Cooperation and Development (OECD) average [1]. Differences in drug costs between the US and other developed countries are particularly evident for cancer drugs, the second largest category of medications sold in the US and the largest proportion (41%) of medications sold worldwide at $37.2 billion USD [2]. For example, Avastin and Erbitux, used for treating metastatic colorectal cancer, are priced at $103,950 and $128,160 per 12-month cycle in the US, while in the UK, they are priced at ~£38,620 and ~£36,154 ($65,635 and $61,445 USD, respectively) [3].

One possible reason for this difference is that the US government does not negotiate with pharmaceutical companies, thus leaving the price subject to market competition. One of the drivers of this policy is the "non-interference" provision in the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2006, which states that the Department of Health and Human Services (HHS) is prohibited from interfering with (price) negotiations for drugs that are covered under the Medicare Prescription Drug Coverage (Part D) [4]. This is unlike many other countries, where governments have enacted caps on drug prices and are able to negotiate prices based on therapeutic benefits. Absence of drug pricing regulation leads to higher drug prices, leaving the burden on taxpayers and businesses.

Another potential mechanism driving drug costs is the approval process of the US Food and Drug Administration (FDA). Indeed, the FDA is relatively efficient in terms of the time for drug review and approval (median 322 days) compared to regulatory agencies in Europe (median 366 days) and Canada (median 393 days) [4]. This is due in part to the Prescription Drug User Fee Act (PDUFA), which was enacted in 1992 and authorized the FDA to collect fees from drug manufacturers. However, despite this apparent efficiency, the FDA's approval process has been questioned on several fronts, including the absence of cost-effectiveness criteria in the process of drug approval [5]. This makes sense, as the FDA's main role is to oversee the safety of drugs, medical devices, food, cosmetics, and other health-related products. However, while the FDA has no role in price setting, it could leverage its position as one of the primary overseers of medications in the US to create another avenue of accountability through the inclusion of cost-effectiveness criteria within the drug approval process.

Incorporating cost-effectiveness criteria into decision-making on healthcare coverage could be accomplished through the Centers for Medicare & Medicaid Services (CMS), which is responsible for health insurance for those with low incomes, the elderly, and the disabled; there are currently around 100 million beneficiaries in the US Indicative of CMS's substantial influence is that commercial payers often follow the lead set by CMS on issues concerning coverage and payment guidelines. Despite the potential for inclusion of cost-effectiveness criteria, one important political driver exists against its use: the Affordable Care Act explicitly prohibits the use of these criteria to make recommendations regarding approved drugs (although not explicitly mentioned in the context of FDA approval processes). More specifically, this is geared towards the Patient-Centered Outcomes Research Institute (PCORI), which was created to fund comparative effectiveness research in which healthcare interventions (e.g., drugs) are compared. This motion was permitted in order to increase access to treatment and put to rest fears of denial of care [6].

ABSENCE OF COST-EFFECTIVENESS CRITERIA

The absence of cost-effectiveness criteria in the drug approval process distinguishes the US from other Western countries, which utilize the Quality Adjusted Life Year (QALY), a measure of disease

Artist: Ryan Xiao
Harvard Medical School
Ryan_Xiao@hms.harvard.edu

burden that takes both the quantity and quality of life lived into account. QALYs are calculated by the amount of time spent in a particular health state weighted by the utility score given to that state; 1 year spent in "perfect health" is equal to 1 QALY, while 1 year spent in "poor health" could be equal to 0.5 QALY. This measure is used in cost-effectiveness analyses (coined "cost-utility" when QALYs are utilized) to allow for a comparison of cost-per-QALY ratios for different interventions or drugs [7]. In the context of the drug approval process, QALYs are a valuation of health benefit resulting from a new treatment or drug.

Internationally, different types of agencies and organizations (with various responsibilities and mandates) involved in healthcare evaluation utilize this measure. However, the QALY thresholds appear somewhat arbitrary and vary substantially by country. In the Netherlands, a maximum threshold value of €80,000 (~$109,500 USD) per QALY gained has been suggested, while in New Zealand and Australia, threshold values of NZ$20,000 (~$17,094 USD) and AU$69,900 (~$65,748 USD) are implicitly used in the drug approval process [8]. In the UK, the National Institute for Health and Care Excellence (NICE) (https://www.nice.org.uk) develops guidance, standards, and information on not only high-quality health care but also social care. Unlike the FDA, NICE is charged with using existing knowledge to inform decisions regarding which treatments should be covered by the UK's publicly funded single-payer system, the National Health Service (NHS). In the case of cost-effectiveness assessment of new drugs, NICE applies a threshold of £20,000–£30,000 (~$33,000–$50,000 USD) per QALY gained. If the QALY-adjusted cost of a new drug exceeds this threshold, there is a substantial risk that the drug will not be approved.

The most important result of the FDA not using cost-effectiveness criteria is that quite a few drugs that are approved in the US (although they might not be covered under public or commercial payers) are rejected in countries such as the UK if they are conveyed to be too expensive with only marginal effects. The 2013 FDA approval of the cancer drugs Kadcyla and Xofigo, for metastatic breast cancer and advanced prostate cancer, respectively, clearly illustrates this effect. While approved by the FDA, the drugs were rejected by NICE based on their cost per QALY gained: Kadcyla and Xofigo currently cost approximately $74,405 and $40,000 per course of treatment, respectively [9,10]. As these new and expensive drugs become preferentially available in the US, they contribute increasingly to the overall high drug prices in the US compared to other Western countries that benefit from value-based pricing methods.

ARE QALYS THE ANSWER?

The combination of increasing drug costs and absence of QALYs in the regulatory process makes it difficult for decision-makers to objectively determine the added value of a drug in such an expansive open market. This complexity and ambiguity is demonstrated by the comparison of 12 cancer drugs approved by the FDA in 2012. Of these, nine were priced at over $10,000 per month, while only three prolonged survival, and two of them by less than two months [11]. While these high costs of care are problematic, the use of QALYs as a standard metric for assessing drug efficacy is not without controversy. One ongoing concern is that the approach discriminates on age and disability. For example, the QALY metric values young and healthy populations more highly than illness-struck, disabled, or elderly populations [6].

Another primary argument against the use of QALYs is its subjective nature. Quality of life varies between individuals and is heterogeneous in meaning, depending on an individual's preferences, circumstances, and experience. The QALY metric assumes that patients value quality of life over length of life, where alternatively, some patients may find more importance in living longer without regard for quality [12]. Additionally, quality of life for an individual does not take into account the individual's impact on caregivers or family. This also entrenches into an ethical dimension of the metric, as it breaches concerns related to subjective conclusions about the benefits and disadvantages of treating different dis-

eases and using different treatment methods, and the evident ethical complications of putting a dollar value on extending a patient's life [12].

CONCLUSIONS

Although there are concerns about the role of QALYs as the sole benchmark of health gains for purposes of allocating resources, decision-making bodies in the US and UK, along with the World Health Organization, have found that they are preferred compared to alternative measures of health improvement [7]. The QALY metric is an important metric to consider in efforts to reduce healthcare costs in the US, as it encompasses both the quality and the quantity of life lived. The QALY allows, in one single metric, a basis with which to compare the effects associated with the use of an intervention. That is, with one metric, there is a means of comparing one treatment against another, or a treatment against no intervention at all. Moreover, this metric is a vehicle for decision-makers to best determine how to allocate scarce resources to maximize health benefits [6,12]. Although the incorporation of QALYs in measuring and comparing health effects will increase the length of the drug review process, it can eventually be a tool to increase due diligence and accountability. By encouraging prices based on real value, drugs could become more affordable and less burdensome to consumers. Although QALYs are heavily debated, there still exists no viable superior alternative method of measuring health benefits in the context of decision making for resource allocation. Moreover, even in case of rejection of the QALY metric, the needs of decision-makers will persist with QALY alternatives that are likely to share many of its attributes.

Although current provisions prevent government-led negotiation of drug prices, if ever approved by Congress, this could be a cornerstone of decreasing drug costs. Additionally, with increasing drug prices and its burden on overall health expenditures, a reform in the drug approval processes and pricing strategies seems imperative. We have highlighted and evaluated just one possibility, introduction of QALYs, but there are potentially many other means. The current emphasis on changing healthcare in the US sets the stage for daring and innovative initiatives to not only improve healthcare but also increase value across the continuum of care. Utilizing healthcare resources in the most optimal manner would benefit all of society.

REFERENCES

1. (2014) OECD Health Statistics 2014: How does the United States compare? Retrieved on June 28, 2015 from http://www.oecd.org/unitedstates/Briefing-Note-UNITED-STATES-2014.pdf.
2. (2014) IMS Health for Healthcare Informatics. Global oncology trend report: Innovation in cancer care and implications for health systems. Plymouth Meeting: The IMS Institute
3. Faden RR, Chalkidou K, Appleby J, Waters HR, Leider JP (2009) Expensive cancer drugs: a comparison between the United States and the United Kingdom. *Milbank Quarterly* 87(4):789-819.
4. Downing NS, Aminawung JA, Shah ND, Braunstein JB, Krumholz HM, Ross JS (2012) Regulatory review of novel therapeutics—comparison of three regulatory agencies. *N Eng J Med* 366(24):2284-2293.
5. Brower A (2008) Is It Time To Take a Harder Look at the QALY? *Biotechnol Healthc* 5(3):47-8.
6. Kirkdale R, Krell J, O'Hanlon Brown C, Tuthil M, Waxman J (2010) The cost of a QALY. *QJM* 103:715-720.
7. Neumann P, Weinstein M (2010) Legislating against Use of Cost-Effectiveness Information. *N Eng J Med* 363:1495-1497.
8. Cleemput I, Neyt M, Thiry N, De Laet C, Leys M (2008) Threshold values for cost-effectiveness in health care Health Technology Assessment (HTA). Brussels: Belgian Health Care Knowledge Centre (KCE); *KCE reports 100C* (D/2008/10.273/96).
9. Palmer E (2014) NICE rejects Roche's hot new breast cancer drug Kadcyla, then invites negotiations. *FiercePharma.* Retrieved on July 3, 2014 from http://www.fiercepharma.com/story/nice-rejects-roches-hot-new-breast-cancer-drug-kadcyla-then-invites-negotia/2014-04-23.
10. Hirschler B (2014) UK cost agency says 'no' to Bayer prostate cancer drug Xofigo. *Reuters.* Retrieved on July 3, 2014 from http://www.reuters.com/article/2014/03/24/us-bayer-britain-idUSBREA2N00120140324.
11. Kantarjian HM, Fojo T, Mathisen M, Zwelling LA (2013) Cancer drugs in the United States: JustumPretium—the just price. *J Clin Oncol* 31(28):3600-3604.
12. Woodward, RM, Menzin, J, Neumann, PJ (2013) Quality-adjusted life years in cancer: pros, cons, and alternatives. *Eur J Cancer Care* 22(1):12-19.

HMSR

VIEWPOINT
Medical Ethics

The "Harm Principle" in the Context of Organ Donation

John Fortunato, MA[1]*

[1]Case Western Reserve University, Cleveland, OH 44106, USA
*Current affiliation: Oakland University William Beaumont School of Medicine, Rochester, MI 48309, USA
Correspondence should be addressed to J.F. (jtfortunato@oakland.edu)

Concern has arisen that "Donation after Circulatory Death" organ donors are not beyond harm; they may feel pain during procurement. A "harm principle," which simply places the prevention of donor harm as a central ethical obligation, should be used during organ donation to protect donor interests and protect donors from physical suffering.

Abstract

Organ procurement in the United States is regulated by the Dead Donor Rule (DDR), which states that individuals cannot be killed by organ donation. Proponents of organ donation have successfully changed death declaration criteria through the creation of brain death and the implementation of shortened "wait times" for the declaration of death via cardiopulmonary criteria, effectively increasing the organ pool. However, substantial concern has arisen that "Donation after Circulatory Determination of Death" (DCDD) donors are not beyond harm; they may feel pain during the procurement process.

Therefore, a "harm principle," which makes the prevention of donor harm a central ethical obligation, should be used during organ donation. A harm principle (1) protects the interests of donors and (2) protects them from physical suffering. If the DDR were revoked and the harm principle was used alone, the pre-mortem donation of both non-vital and vital organs could be permitted (under strict guidelines formulated to protect the donor). However, it is more realistic that the harm principle be implemented in conjunction with the DDR. This would prohibit DCDD donation without the use of anesthesia. It could also permit the pre-mortem procurement of non-vital organs from those who are imminently dying. Through this analysis, the implications of the harm principle on different methods of organ procurement will be examined.

Introduction

The DDR prohibits the killing of individuals by or for organ donation [1]. In other words, potential organ donors must be dead before their organs may be procured for donation (excluding the living donation of non-vital organs, such as a kidney). However, the utilitarian appeal of an increased organ supply has lead to the reconstruction of death declaration criteria, first by the recognition of brain death [2], and then by the formation of DCDD protocols. Since the conception of DCDD protocols, substantial concern has arisen over the validity of death declarations in the cases of DCDD donors [3-5]. In fact, DCDD donors may feel pain during the organ procurement process [6-9]. Such evidence begs the question: "Are these donors really dead?" Using a term coined by Miller, Troug, and Brock, our adherence to the DDR may be nothing more than a "moral fiction," used only to satisfy public opinion of the medical field: that doctors must not intentionally cause the death of their patients [10].

Proponents of organ donation have successfully circumvented the DDR through a conceptual gerrymandering of death declaration criteria, leaving donors unprotected and subject to substantial harm in the donation process. Therefore, a "harm principle" should be developed as an ethical standard for organ transplantation services. A nearly identical proposal was made by Rodriguez-Arias *et al*.: "what is important for the protection and respect of potential donors is […] to be certain that they are beyond suffering and to guarantee that their autonomy is respected" [11]. It is the goal of this paper to expand upon this proposal, dissecting the practical consequences that a "harm principle" would have on the different methods of organ procurement. Finally, a practical approach, involving the application of the harm principle in conjunction with the already used DDR, will be explored.

THE ACCEPTANCE OF BRAIN DEATH

Brain death criteria were first proposed by the Ad Hoc Committee at Harvard Medical School in 1968 [2]. Following its "rollout," the Harvard criteria were rapidly accepted as a valid benchmark for the declaration of death in the United States. In 1981, the President's commission recognized death of the entire brain as death. Today, all 50 states have laws that recognize brain death as death [12].

Despite its initial rapid growth and acceptance, several scholars have contested the validity of brain death criteria as a true measurement of death. In 1998, Shewmon wrote the seminal paper discrediting the validity of brain death as death. While some scholars, such as James Bernat, have made philosophical arguments supporting brain death, arguing that it is an appropriate indicator of the permanent cessation of functioning of the organism as a whole [13], Shewmon contests such arguments. Instead, he contends that the bodies of brain dead patients can still perform some integrated functioning (maintain homeostasis, grow, experience puberty, gestate a fetus), sometimes for many years [14]. Other scholars, like Youngner and Truog, have echoed similar concerns [15,16].

It should be noted that the validity of brain death criteria for declaration of death is of no consequence to the arguments made in this paper. Brain dead patients are believed to be beyond harm. Therefore, so long as their previous informed consent had been obtained, the harm principle would allow for donation of organs from the brain dead (regardless of their legal status as dead or alive). Still, the intent with which brain death was created is of great importance. The Harvard committee directly cited two reasons for the formation of brain death: (1) to avoid homicide charges against physicians who turned off the ventilator of brain dead patients and (2) to avoid similar charges for the procurement of their organs [2]. No philosophical arguments were made as to why brain death is death. Instead of intending to more precisely identify death, these criteria were written with a utilitarian motive: to increase the donor pool of organs. Not coincidentally, about 90% of the organs procured are from brain dead patients [17]. It was this utilitarian intent and a loss of focus on donor protection that opened the door to donor harm in the organ procurement process.

DCDD PROTOCOLS

These same motives inspired the formation of DCDD protocols. Key to the ethical debate over DCDD protocols is the "waiting time" required after cessation of cardiopulmonary functioning for the patient to be declared dead. How long must you wait after cessation of cardiopulmonary functioning until that cessation is irreversible? Some countries in Europe wait for 10 minutes of asystole and apnea [18]. The Institute of Medicine recommends 5 minutes [19]. The University of Pittsburgh requires just 2 minutes [20]. Finally, Denver Children's Hospital requires only 75 seconds for some neonatal heart donors [21].

Irreversibility, while intrinsic within the definition of death, is difficult to define. There has been much disagreement on how much time must elapse for cardiopulmonary functioning to be irreversibly lost, as evidenced by the wide range of waiting times involved in different protocols. However, a recent re-

search study has reported the successful resuscitation of animals after 11 minutes of asystole [22]. Such findings cast doubt on the validity of protocols with shorter waiting times. Is the "death" of DCDD donors after 5 minutes, 2 minutes, or 75 seconds really irreversible?

Proponents of DCDD donation would contest this question using an argument formulated by James Bernat: DCDD donors do not need to be irreversibly dead (cannot be resuscitated). They must only be permanently dead (will not be resuscitated), as permanence is an acceptable indicator of pending irreversibility [23]. Permanent loss of cardiopulmonary functioning occurs when there is no chance for spontaneous auto-resuscitation (about 7 minutes) [24], much sooner than the irreversible loss of those same functions. Bernat justifies the use of "permanence" instead of "irreversible" in death declaration only by citing common medical practice: that is how death is declared in US hospitals today [7].

Just like the formation of brain death criteria, the criteria used for diagnosing death in DCDD donors has no philosophical basis. It was formulated for practical convenience: the permanence standard is what doctors currently use. In addition, it conveniently satisfies utilitarian motives to increase the donor pool: "permanence" permits a shorter wait time after cessation of cardiopulmonary functioning. This limits organ damage from warm ischemia that would otherwise ensue and allows healthier organs to enter the organ pool.

Although DCDD protocols work to ensure the procurement of healthy organs for donation, concern for the welfare of DCDD donors during the donation process is lacking. As a result, DCDD donors may inadvertently experience pain during organ procurement.

Many consider cardiopulmonary criteria a valid tool for death declaration because it acts as a surrogate indicator of the impending brain death that will follow shortly after [13]. However, in cases of DCDD organ donation, no brain tests are carried out to confirm brain death. If organs are procured immediately after permanent cessation of cardiopulmonary function, the brain may not yet be "dead," and that person may feel pain.

These concerns are more prominent when extracorporeal membrane oxygenation (ECMO) is used. Often, after patients are declared dead, they are placed on ECMO, which circulates oxygenated blood throughout their body. This is done to keep their organs "fresh" and to eliminate the need for immediate organ retrieval [25]. However, just as oxygenated blood travels to their soon-to-be donated organs, so too does it travel to the brain. In this way, ECMO may restore brain functioning if it had only been permanently, but not irreversibly, lost. If brain function is restored, there is a very real chance that pain is experienced through the procurement process.

As a response, anesthetic could be used as a safeguard for DCDD donors, especially for those on ECMO. An inflated thoracic occlusion balloon has also been used in conjunction with ECMO to block oxygenated blood flow to the brain (and thus eliminate the potential for brain revival) but allow blood flow to the rest of the body [26]. These practices have met some resistance. Bernat and others have rejected the use of both a thoracic occlusion balloon and ECMO for donors, as the former raises questions of physician complicity in the casuistry of death, and the latter negates the determination of death by violating the permanence standard [6,7]. Such practices could also result in general public distrust of the medical profession and organ donation in general.

Not surprisingly, shortened DCDD criteria have created many practical dilemmas and invited many scholarly critiques. The notion of permanence has allowed many to falsely justify the premature declaration of death of organ donors, only to negate that death declaration by placing patients on ECMO and making the "permanent" loss of circulation no longer permanent. This is precisely why the use of permanence instead of irreversibility is suspect: it allows the intentions of others to play a role in de-

termining death. In these cases, death declaration is contorted as a means to the utilitarian end of organ procurement.

Further, these practices have opened Pandora's box, challenging common sense understandings of death, by creating logically counterintuitive situations where one might cause physical harm to the "dead." The use of "permanence" may be ethically preferable in most cases, where death declaration does not have surgically invasive and potentially harmful consequences. But it is only in situations of consequence (organ donation) when we are completely dependent on the accuracy (and thus validity) of our death declaration criteria. Ironically, it is only in these circumstances, when it really matters, where such a construal of permanence falls short.

THE DDR AND ITS PRACTICAL INEFFECTIVENESS

These debates have blurred the lines between life and death, with little hope for clear consensus in the near future. The criterion used for death declaration is dependent on one's philosophical definition of death. In our pluralistic society, such definitions are subjective to the individual and are socially constructed [15]. Thus, comprehensive social agreements on the definition of death, death declaration criteria, and related policy are unlikely to be reached.

Still, any definition of death would prove to be incompatible with the notion of "post-mortem physical pain and suffering." DCDD donors, if they are capable of feeling pain, are surely not dead, at least in the way we commonly understand death. Their "death," at the point of donation, is merely a legal formality. In practice then, the DDR does not protect donors from a painful donation. Conceptual gerrymandering of death criteria has stripped the DDR of its ability to do so. Instead, it only promotes a sense of public trust of the medical profession, and it allows society (including physicians) to believe a "moral fiction": that organs are only procured from people after they are dead [8].

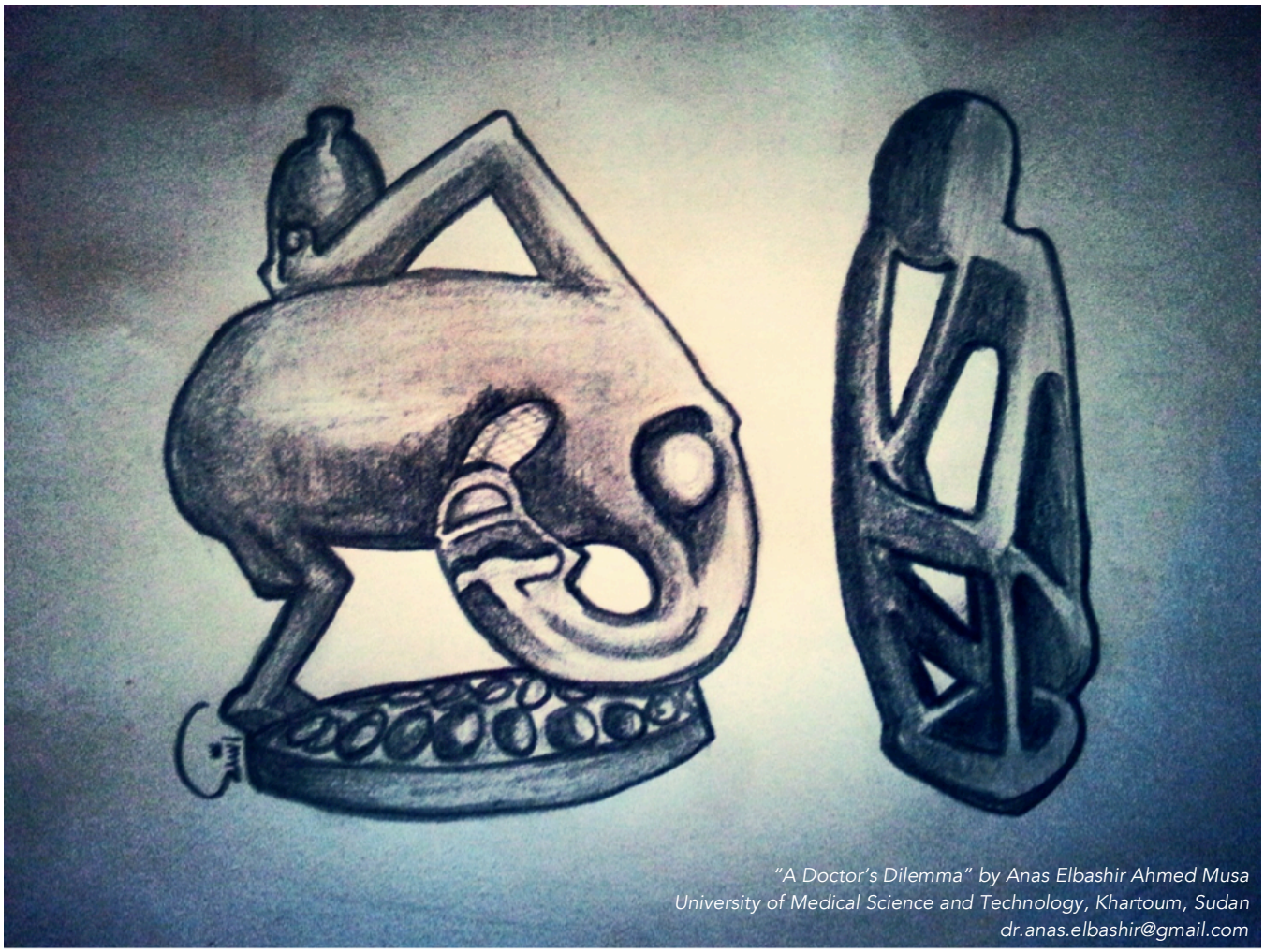
THE NEED FOR A NEW ETHICAL GUIDELINE: THE HARM PRINCIPLE

To protect organ donors, a new ethical standard must be created that does not engage in already exhausted, futile arguments over the validity of death criteria. While the DDR was formulated to protect organ donors, pro-organ donation utilitarianism has restructured death criteria and has left the DDR ineffective. Organ procurement services ought to follow the harm principle, which bypasses the death declaration criteria debate and is thus immune from similar gerrymandering. The harm principle simply states: "organ donors must not be harmed by the organ procurement process."

Harm has been defined as simply "a setback to an interest" [27]. In other words, if the interests of an individual have been impeded, then that individual has been harmed. The protection of patient interests has been a long-standing priority in medical practice. The interests of patients have been protected through the patient's ability to formulate living wills and advance directives through requirements for informed consent and through the overall promotion of patient autonomy, the chief principle of American bioethics.

In this way, individuals define their own interests. They effectively decide for themselves what is, and what is not, harm. However, in this context of the harm principle for organ donation, patient autonomy does have a limit. The harm principle does not permit for a physically painful donation process, even if informed consent were obtained. While individuals may possess the right to inflict physical harm upon themselves, physicians certainly retain the right to refrain from complicity in physical harm of their patient.

The right of physician integrity has gained foothold in the futility debate. If patient care is determined to be medically futile, physicians have the right to refrain from offering that care [28]. In fact, they ought not offer futile care to patients, for such actions are contrary to the goals of the medical profession. Physician integrity, in these cases, justifi-



"A Doctor's Dilemma" by Anas Elbashir Ahmed Musa
University of Medical Science and Technology, Khartoum, Sudan
dr.anas.elbashir@gmail.com

ably trumps patient autonomy. Along a similar vein, the explicit physical harm of organ donors directly opposes the goals of the medical profession, even if the donor knowingly consents to it. Physical harm can be easily prevented by anesthesia, and it should always be avoided. Therefore, the harm principle in the context of organ donation is twofold: (1) organ procurement must not interfere with the interests of the organ donor, and (2) organ procurement must not inflict physical suffering onto the organ donor.

The harm principle must be explored in two separate scenarios: when used in conjunction with the DDR, and when used alone, after the revocation of the DDR. While it has been shown that the DDR today serves little functional purpose in the protection of donors, it is unlikely to be revoked in the near future. Although the harm principle would be adequate if it stood alone, the more likely scenario, where it is implemented alongside the DDR, is a sufficient, "better than nothing" compromise. Under this approach, both the DDR and the harm principle must be satisfied for donation to be ethically acceptable.

THE HARM PRINCIPLE USED IN CONJUNCTION WITH THE DDR

Donation after Declaration of Brain Death

Brain dead patients, in current medical practice and under current criteria, are understood to be dead. Therefore, the donation of organs from brain dead patients is not a violation of the DDR.

Although the validity of brain death has been debated, it is unanimously agreed upon that brain dead patients are beyond harm. Therefore, donation after declaration of brain death would be unaltered

by the use of a harm principle in conjunction with the DDR.

DCDD With or Without ECMO

The use of a harm principle would have its largest impact on the present-day DCDD practices. Obviously, DCDD protocols are in compliance with death declaration criteria and thus "follow" the DDR. However, substantial likelihood for harm, discussed in detail above, must be addressed.

DCDD, with or without the use of ECMO, often utilizes the permanence standard. As shown above, there is a real possibility that DCDD donors are not yet brain dead and could feel their organ procurement operation. When ECMO is used, there is an even greater chance for brain functioning to be regenerated and for pain to be experienced. Here, the second component of the harm principle is violated: physical pain may be inflicted onto the organ donor.

The first part of the harm principle is also violated: DCDD, with or without ECMO, interferes with the interests of the donor. Organ donors consent to donation after death. Most, if not all, organ donors would reject a notion of death in which you are still "alive enough" to feel pain. Truly informed consent was never obtained; organ donors consent to organ donation under the presumption that they will be dead, and thus intrinsically immune to physical pain, during the procurement process. Due to the violations of both components of the harm principle, DCDD performed in these circumstances should be eliminated.

DCDD could still be permitted circumstances where proper measurements are taken to promote the interests of donors and protect them from physical pain. One option would be to wait a longer duration after the loss of cardiopulmonary functioning to declare death. The permanence standard allows death to be declared in close proximity to the loss of breathing and circulation, before brain death has occurred, thus opening the door to potential harm of the donor. By simply waiting longer (10 minutes or more), physicians could ensure the irreversible death of brain function, which would eliminate the possibility of the donor perceiving pain during the procurement process.

Although this practice would be ethically justified by the harm principle, it may be impractical without the use of ECMO. Ten minutes of warm ischemic time before procurement may damage the viability of organs. Studies have shown, however, that organs still retain their viability if the donor body is placed on ECMO after 10 minutes or more of pulselessness [26]. This would be ethically acceptable, because the brain would be irreversibly dead at this point, and could not be resuscitated by the perfusion of oxygenated blood. Therefore, a waiting time of 10 minutes after the commencement of pulselessness would require the use of ECMO to keep the organs viable for donation.

ECMO could also be used earlier, after the declaration of "permanent" death. However, other measures must be taken to ensure donor protection from physical suffering. The use of anesthesia would not prevent oxygenated blood flow from going to the brain, but it would successfully eliminate the chances of donor perception or awareness, just as it does during surgery for living patients. This would eliminate the possibility of donors perceiving pain during organ procurement. This practice may be preferable, as it prevents physical harm, ensures less warm ischemic time and higher viability of the organs, and is more practical than waiting 10 minutes.

The use of a thoracic occlusion balloon for DCDD donors must be rejected. Using the permanence standard, the placement of an occlusion balloon would be invasive and could case pain and suffering (similar to organ procurement during permanent death). Therefore, it does not satisfy the harm principle. If the procedure were performed in conjunction with anesthesia, the balloon would serve no purpose, as anesthesia already adequately protects the donor from physical harm. Therefore, the placement of an occlusion balloon would be unnecessarily invasive and without benefit.

As with all medical operations, the consent obtained for these procurement processes must be truly informed. Patients and their families ought to be made aware of the use of anesthetic during the procurement process. Concerns that donors may refuse to donate, knowing that they will need anesthesia, does not justify the hiding of relevant medical information. As with all medical decisions, it may be impractical for surrogates to be completely informed in the short time after death and before DCDD donation. Still, efforts should be made to educate surrogates, to educate the public in general, and to educate individuals when they declare themselves as organ donors.

In summary, DCDD done with or without ECMO is justified under the practical interpretation of the DDR, but is unjustified under the harm principle. These practices commonly use the "permanence" standard, and procurement may begin before brain death, when the donor can still perceive physical pain. In order to operate in compliance with the harm principle, several options could be exercised. To eliminate physical pain, the wait time for procurement could be extended to 10 minutes. To preserve the organs of the donor, ECMO may be used after the 10-minute wait time, when the brain is irreversibly dead and the donor cannot be harmed. Alternatively, ECMO could be used after the declaration of "permanent" death. Because the brain has not irreversibly died, anesthesia must be used to restrict the donor's ability to feel physical pain. As with all organ donations, a stringent informed consent process is required in order to protect and promote the interests of donors.

Pre-Mortem Donation of Non-Vital Organs

The DDR, while primarily arguing that individuals cannot be killed by organ donation, can also be interpreted in its most literal sense: organs cannot be procured from individuals until they die [1]. Despite this rule, it is common practice for kidneys to be procured from living donors. Thus, our society has made a conscious exception to the DDR.

If donors need not be dead for the procurement of their organs to be justifiable, then what prevents us from procuring non-vital organs from consenting organ donors before withdrawal of their life-sustaining care? Discussions of organ donation could not begin until the decision to withdraw care had already been independently made. Youngner and Arnold first proposed such an argument, insisting that a clear and rigorous consent process accompany it. More recent arguments have also been made for kidney donation before end of life care [29]. In these cases, advance directives could dictate, "If a medical decision is made to withdraw care, resulting in my death, I wish to first donate both of my kidneys." That patient could be taken to the OR in the morning, undergo a kidney procurement operation under normal anesthesia, and have their bleeding vessels tied off or cauterized. Later in the day, the decision to withdraw care could be exercised as planned. The patient would die as a result of the care withdrawal, far before he/she could die from renal failure [1].

These actions are in accordance with the harm principle. First, given a stringent consent process, they act in accordance with the interests of the patient. Second, normal pain management measures would be taken, just as they are already taken both with surgical patients and with patients prior to the withdrawal of care. Therefore, the patient would be protected from unwarranted physical harm.

Just as with living kidney donation today, premortem donation would require that an exception be made to one construal of the DDR: that organs cannot be procured from patients until they die. However, the more stringently followed component of the DDR, that individuals cannot be killed by organ donation, is still followed. Therefore, these practices seem to operate within some confines of the DDR. The procurement of organs would not lead to the death of these organ donors. Rather, the independently determined withdrawal of life-sustaining treatment more proximately causes the death of these patients.

The pre-mortem donation of non-vital organs provides us with a new, ethically defensible way to increase the organ pool without causing harm to or-

gan donors. Further, it promotes the autonomy of dying patients, giving them the opportunity to offer life to others through an organ donation process that is free of harm.

THE HARM PRINCIPLE IF THE DDR WERE ELIMINATED

Due to the DDR's current inability to adequately protect donors from harm, some have argued for the elimination of the DDR from the ethical framework that governs the organ procurement processes [8,9,11]. If the harm principle alone were used to oversee organ donation, all of the procurement options and guidelines discussed above would remain acceptable. However, additional options for organ procurement would become available.

Pre-Mortem Donation of Vital and Non-Vital Organs

A controversial proposal to provide patients with the option to donate all of their organs while still alive (yet imminently dying) has been lead by Franklin G. Miller and Robert D. Truog [8]. These scholars build their argument on the foundational claim that withdrawing life support is not merely "allowing the patient to die," but is in fact a justifiable "killing" of the patient. Dan Brock has also made coinciding arguments, stating, "The distinction between "killing" and "allowing to die" is conceptually confused and mistaken" [1,30]. If the withdrawal of care is a justifiable killing of patients (as medical practice indicates), Miller and Truog contend that so too could be the killing of patients via organ donation (in restricted cases under a strict consent process). Frankly, many patients consent to organ donation but, after withdrawal of care, do not die quickly enough for their organs to retain viability and be donated. In these circumstances, a prohibition of pre-mortem organ donation restricts the dying patient's ability to have their wishes to donate their organs be expressed. Pre-mortem donation of vital organs would promote patient autonomy by providing these patients with an option to die painlessly from consensual organ donation, rather than from the sometimes drawn-out process associated with the withdrawal of life-sustaining treatment.

As a safeguard to the utilitarian killing of individuals for their organs, discussions to donate pre-mortem could only begin after the decision to withdraw life-sustaining treatment were independently agreed upon by the patient and the health care team. After that decision is made, the patient could be approached with the option to die via organ donation, rather than die via withdrawal of care. In these cases, patients would be taken into the operating room, and all of their procurable organs would be removed under full anesthesia. The procurement of vital organs (the heart and lungs) would naturally result in the death of the donor. If these organs were not donated, then a post-procurement withdrawal of life-sustaining treatment would directly lead to the patient's death.

The use of a stringent informed consent process (requiring informed patient consent, living will authorization, or informed surrogate consent), along with the administration of anesthesia, would satisfy the first and second components of the harm principle, respectively. In addition, these protocols would provide dying patients with increased autonomy in the end of their lives, giving them the altruistic option for organ donation that would, under the DDR, be otherwise forbidden.

Pre-Mortem Donation of Vital and Non-Vital Organs for Persistent Vegetative State Patients

Just as ICU patients could choose to die from organ donation rather than from the withdrawal of life sustaining treatment, persistent vegetative state (PVS) patients could also be provided with this option. As seen in the famous Terri Schiavo case, the decision to withdrawal life-sustaining treatment from PVS patients can result in a drawn-out starvation process, slowly leading to an eventual death. Rather than restricting the inevitable death of PVS patients to starvation and dehydration, the pre-mortem donation of organs offers a quicker and more humane death.

In contrast to conscious patients who would choose pre-mortem donation, PVS patients have no current interests to protect. But, as with other unconscious patients, the previously expressed interests of these patients (expressed through a living will or through surrogate decision-makers) must be respected. In addition, anesthesia ought to be used as a safeguard to prevent even the remote chance of physical harm. These provisions would satisfy the harm principle, adequately protecting organ donors.

A Uniform Procurement Process

It should be noted that the elimination of the DDR would drastically simplify organ donation processes across the board. Without the need to declare death, all donors could be brought to the operating room with active cardiopulmonary function. Since organ procurement is a justifiable cause of death in some cases, the same procurement procedure that is commonly used for brain dead donors could be used universally. Anesthesia would be required for all patients who are not beyond harm (all non-brain dead donors). While procurement does not affect the declaration of death for brain dead donors, it would effectively lead to the death of still-alive donors. Therefore, although the harm principle permits some DCDD donations, such practices are impractical without the requirements of the DDR. It would be gratuitous for organ procurement teams to withdraw the life sustaining treatment of DCDD donors, wait for their "permanent" death, and then re-start life sustaining treatment in preparation for procurement. Such processes would probably no longer be used if the DDR were revoked.

MOVING FORWARD IN OUR CULTURAL CONTEXT

Due to the overpowering utilitarian appeal of an increased organ donor pool, the criteria used to declare death have been altered to permit the procurement of organs earlier in the dying process. These practices have opened the door to potentially harmful organ procurement operations, where both the interests and physical well being of organ donors are jeopardized. However, to satisfy the DDR, these patients are still officially declared dead. It would be ethically ideal to revoke the DDR, and use only the harm principle to regulate organ donation. The primary role of ethical guidelines in organ procurement is the protection of the donor. The harm principle, used without the DDR, completely fulfills this role. Further, revocation of the DDR provides the medical profession with an opportunity to "come clean" with the public, instead of operating under the "moral and legal fiction" that organ donors are always dead when their organs are procured.

However, removal of the DDR and admittance of its fabrications could create serious public distrust of the medical profession, leading to a general reluctance to donate organs. Politically, the revocation of the DDR would be characterized as an infringement on one's right to life, and would lack initial support. In addition, guidelines dictated by the harm principle are crude and need to be further developed. Pre-mortem donation of vital organs, once publicly known and accepted, could be used by patients as a motive to pre-maturely justify their own determination to withdraw life-sustaining treatment. In other words, individuals may consent to the withdrawal of life-sustaining treatment sooner in their course of care than they otherwise would have, if there were no prospect for pre-mortem donation. Thus, safeguards need to be developed to increase the usefulness of the harm principle.

Although the DDR is imperfect and fosters public deception, its abolition seems unlikely, given the cultural shift required for its acceptance. Therefore, use of the DDR and the harm principle in conjunction would be ideal. Even with the DDR in place, the harm principle is able to additionally protect the interests and prevent the physical harm of organ donors, primarily by limiting DCDD donation. Further, its promotion of patient interests provides organ donors with more autonomy at the end of life. While this approach is not without limitations, it may be the most practically and politically feasible, offering adequate and expedient protection from harm for organ donors. Until public acknowledgement of the flawed death declaration criteria and the inadequacies of the DDR become widespread,

this proposal will, for now, serve as an adequate "halfway" approach, ensuring the protection of organ donors from harm.

REFERENCES

1. Arnold RM, Youngner, SJ (1993) The dead donor rule: should we stretch it, bend it, or abandon it? *Kennedy Inst Ethics J* 3:263–278.
2. (1984) Landmark article Aug 5, 1968: A definition of irreversible coma. Report of the Ad Hoc Committee of the Harvard Medical School to examine the definition of brain death. *JAMA* 252:677–679.
3. Harrington MM (2009) The Thin Flat Line: Redefining Who Is Legally Dead in Organ Donation After Cardiac Death. *Denver Univ Law Rev* 86:335–380.
4. Marquis D (2010) Are DCD Donors Dead? *Hastings Cent Rep* 40:24–31.
5. Joffe AR, Carcillo J, Anton N, deCaen A, Han YY, Bell MJ, Maffei FA, Sullivan J, Thomas J, Garcia-Guerra G (2011) Donation after cardiocirculatory death: a call for a moratorium pending full public disclosure and fully informed consent. *Philos Ethics Humanit Med* 6:17.
6. Bernat JL, Capron AM, Bleck TP, Blosser S, Bratton SL, Childress JF, DeVita MA, Fulda GJ, Gries CJ, Mathur M, et al. (2010) The circulatory-respiratory determination of death in organ donation. *Crit Care Med* 38:963–970.
7. Bernat JL, Bleck TP, Blosser SA, Bratton SL, Capron AM, Cornell D, Devita MA, Fulda GJ, Glazier AK, Gries CJ, et al. (2013) Circulatory Death Determination in Uncontrolled Organ Donors: A Panel Viewpoint. *Ann Emerg Med* 63:384-90.
8. Bernat JL (2013) On Noncongruence between the Concept and Determination of Death. *Hastings Cent Rep* 43:25–33.
9. Miller FG, Truog RD (2011) *Death, Dying, and Organ Transplantation: Reconstructing Medical Ethics at the End of Life* (Oxford University Press).
10. Miller FG, Truog RD, Brock DW (2010a). The Dead Donor Rule: Can It Withstand Critical Scrutiny? *J Med Philos* 35:299-312.
11. Miller FG, Truog RD, Brock DW (2010b) Moral Fictions and Medical Ethics. *Bioethics* 24:453-60.
12. Rodriguez-Arias D, Smith MJ, Lazar NM (2011) Donation After Circulatory Death: Burying the Dead Donor Rule. *Am J Bioeth* 11:36–43.
13. (1981) United States President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. *Defining death: a report on the medical, legal and ethical issues in the determination of death.*
14. Bernat JL, Culver CM, Gert B (1981) On the Definition and Criterion of Death. *Ann Intern Med* 94:389-94.
15. Shewmon DA (1998) Chronic "brain death": meta-analysis and conceptual consequences. *Neurology* 51:1538–45.
16. Youngner SJ, Arnold RM (2001) Philosophical Debates About the Definition of Death: Who Cares? *J Med Philos* 26:527–537.
17. Truog RD (1997) Is it time to abandon brain death? *Hastings Cent Rep* 27:29–37.
18. Bresnahan MJ, Mahler K (2010) Ethical debate over organ donation in the context of brain death. *Bioethics* 24:54–60.
19. Dhanani S, Hornby L, Ward R, Shemie S (2012) Variability in the determination of death after cardiac arrest: a review of guidelines and statements. *J Intensive Care Med* 27:238–252.
20. DeVita MA, Snyder JV (1993) Development of the University of Pittsburgh Medical Center policy for the care of terminally ill patients who may become organ donors after death following the removal of life support. *Kennedy Inst Ethics J* 3:131–143.
21. Boucek MM, Mashburn C, Dunn SM, Frizell R, Edwards L, Pietra B, Campbell D (2008) Pediatric Heart Transplantation after Declaration of Cardiocirculatory Death. *N Eng J Med* 359:709–714.
22. Stiegler P, Sereinigg M, Puntschart A, Seifert-Held T, Zmugg G, Wiederstein-Grasser I, Marte W, Meinitzer A, Stojakovic T, Zink M, et al. (2012). A 10 min "no-touch" time - Is it enough in DCD? A DCD Animal Study. *Transpl Int* 25:481–492.
23. Bernat JL (2010a) How the Distinction between "Irreversible" and "Permanent" Illuminates Circulatory–Respiratory Death Determination. *J Med Philos* 35:242–255.
24. Bernat JL (2010b) How autoresuscitation impacts death determination in organ donors. *Crit Care Med* 38:1377–1378.
25. Ko WJ, Chen YS, Tsai PR, Lee PH (2000) Extracorporeal membrane oxygenation support of donor abdominal organs in non-heart-beating donors. *Clin Transplant* 14:152–156.
26. Magliocca JF, Magee JC, Rowe SA, Gravel MT, Chenault RH, 2nd, Merion RM, Punch JD, Bartlett RH, Hemmila MR (2005) Extracorporeal support for organ donation after cardiac death effectively expands the donor pool. *J Trauma* 58:1095–1101; discussion 1101–1102.
27. Feinberg J (1984) *Harm to Others* (Oxford University Press).
28. Schneiderman LJ, Jecker NS, Jonsen AR (1990) Medical futility: its meaning and ethical implications. *Ann Intern Med* 112:949–954.
29. Morrissey PE (2012) The case for kidney donation before end-of-life care. *Am J Bioeth* 12:1–8.
30. Brock DW (1992) Voluntary Active Euthanasia. *Hastings Cent Rep* 22:10–22.