Risk-Based Strategies for Population Screening and Disease Management

Citation

Permanent link

Terms of Use

Share Your Story

Risk-based Strategies for Population Screening and Disease Management

A dissertation presented

By

Vidit Nikhil Munshi

To

The Committee on Higher Degrees in Health Policy

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

In the subject of

Health Policy

Harvard University

Cambridge, Massachusetts

May 2020

Risk-based Strategies for Population Screening and Disease Management

Dissertation Abstract

The objective of this dissertation is to explore the tradeoffs of risk-based strategies to screen populations for disease and manage patients with abnormal test outcomes. Many guidelines on screening or management of a patient population are based on a "one-size-fits-all" approach. However, tailoring guidelines to sub-groups based on risk has the potential to improve efficiency in health care spending and can result in adoption of strategies which decrease health care costs and improve patient outcomes

Chapter 1 introduces a risk-stratified approach to screening for cervical and colorectal cancer. Published literature indicates that early negative screens may predict decreased future risk of cancer incidence. This suggests that individuals with negative screens early on may not be harmed by extending their screening interval moving forward, potentially presenting a significant cost-savings to the health care system. In this chapter, we use simulation modeling to explore the cost-effectiveness of strategies to widen the screening interval for women with early single or serial negative screens for both cervical and colorectal cancer. For cervical cancer, we find that an adaptive strategy to extend the screening interval from five years up to 15 years after just 1 negative screen is cost-effective compared to current guideline screening. On the other hand, guideline screening using fecal immunochemical testing (FIT) for colorectal cancer is cost-effective compared to any adaptive strategy.

Chapter 2 evaluates the cost-effectiveness of updated guidelines by the American Society for Colposcopy and Cervical Pathology (ASCCP) on the management of abnormal cervical cancer screening results. While previous guidelines in 2012 were based on specific actions for specific test results, updated guidelines in 2019 have shifted to a risk-based approach with the goal of applying "equal management of equal risks." We modified a microsimulation model of cervical cancer to assess the cost-effectiveness and resource utilization associated with newer guidelines compared to previous guidelines. We find that, under current screening practices, the 2019 guidelines are cost-effective and cost-saving relative to previous guidelines, indicating that a risk-based approach improves the efficiency of cervical cancer screening and management.

Chapter 3 explores diagnosis and screening for post-transplantation diabetes mellitus (PTDM), a complication of solid organ transplantation. There is sparse and inconsistent literature on PTDM, leading to inconsistent estimates of incidence and no differentiating guidelines on how to manage transplantation patients who may be at risk for PTDM. We analyzed data sets from kidney, liver, and heart transplantation patients at the Mayo Clinic, and used data imputation and simulation modeling to evaluate the potential impact of a screening program to collect hemoglobin A1c (HbA1c) and fasting blood glucose (FBG) from all patients in the immediate post-transplantation setting. We find that poor and inconsistent collection of HbA1c and FBG results in underestimating of PTDM incidence and that better screening of these measures may be a cost-effective intervention to improve long-term patient outcomes.

Table of Contents

Acknowledgements

This work would not have been possible without the teamwork of many terrific researchers. Thanks to the cervical cancer team at CHDS for letting me be a part of it, and to the team at the Mayo Clinic for all your hard work.

Thanks to my outstanding committee – Jane, Soroush, and Ankur – for your guidance, patience, and friendship. You have made me a better researcher, teacher, and mentor.

Thank you to my incredible family and friends for all of your support through the years – I couldn't have done it without you.

This dissertation is dedicated to my loving wife and son, Nisha and Liyan. WE did it!

CHAPTER 1

Cost-effectiveness of A Risk-stratified Approach to Improve Cancer Screening Efficiency

ABSTRACT

BACKGROUND

US Preventive Services Task Force (USPSTF) guidelines for cervical and colorectal cancer screening do not consider changes in interval based on pervious screen results. Recent literature suggests early negative screens may be predictive of reduced downstream risk. We explored a risk-stratified approach to cancer screening in which one or multiple negative screens extend the screening interval and improve efficiency of cancer screening.

METHODS

We utilized two microsimulation models which have been used in previous studies to inform USPSTF guidelines. We analyze multiple scenarios of negative screens needed to extend screening to various lengths on USPSTF guidelines for cytology-based and co-test screening for cervical cancer and fecal immunochemical test (FIT) screening for colorectal cancer. The number of screens, cancer cases, and deaths were estimated for each scenario and a cost-effectiveness analysis was conducted to determine the optimal strategy.

RESULTS

An adaptive screening strategy that extended co-test screening to 15 years after 1 negative screen was the cost-effective strategy (ICER $62,700/QALY) for cervical cancer screening. For colorectal cancer screening, guidelines annual FIT screening was cost-effective (ICER $70,300/QALY). Cervical cancer adaptive screening strategies significantly reduced the number of screens with minimal impact on life expectancy.

CONCLUSION

A risk-based adaptive screening approach can improve efficiency in screening by accounting for patient heterogeneity in the screening population identified by prior screening results. Considering adaptive, personalized screening strategies in future guideline development for cancer screening may help to reduce screening costs and fund other cost-effective interventions without placing a burden on the health care system.

BACKGROUND

The U.S. Preventive Services Task Force (USPSTF) currently recommends screening for two cancers with an "A" rating, indicating high certainty of substantial benefit (1, 2). Cervical cancer screening using cytology (Pap) testing is recommended every 3 years for women aged 21-29 years. From ages 30-65 years, women have the option of either continuing cytology every 3 years, switching to HPV testing alone every 5 years, or switching to cytology in combination with HPV testing (i.e., "co-testing") every 5 years. Additionally, colorectal cancer screening is recommended with the option of using several different tests, such as colonoscopy or fecal immunochemical tests (FITs) at different time intervals for adults aged 50-75 years. One common theme across all recommended cancer screening protocols is that the re-screening interval between screens is constant across the duration of screening irrespective of screening history, which may stem from the clinical trials and modeling studies that informed these recommendations using a constant re-screening interval.

As data have been gathered on implemented screening programs over the past decade, there is growing evidence that negative screen results can provide actionable information. Women who both have a normal cytology result and test negative for a high-risk HPV infection (i.e., co-test negative) have a substantially reduced risk of cervical cancer for at least five years compared to those with either a positive cytology or a positive HPV test result (3-5). Additionally, an observational cohort study found that lengthening the interval between screens to beyond five years may be "feasible and safe (6)" for those with a co-test negative result. Similarly, an analysis of patients who have undergone colonoscopy screening for colorectal cancer found that colorectal cancer risk is reduced for longer than 10 years following a negative

screen (7). While a modeling study for the USPSTF found that colonoscopy screening could be extended from 10 to 15 years if screening began at age 45 (8), no study has evaluated the potential for reducing screening on the basis of one or multiple negative screens. Extending the interval between screening tests can drastically reduce the burden and the harms associated with screening, including risk and disutility associated with false-positive results, as well as monetary costs, which have been estimated to be in the billions of dollars for screening and follow-up (9). Whether early negative screen results are enough of an indicator of low risk to warrant screen interval extension and the potential loss of health benefits is uncertain.

Decision models provide a useful tool in simulating data and projecting outcomes, particularly in cases where the time horizon of available data is limited, and clinical trials do not exist to test specific strategies. In recent years, decision models have been used to aid in clinical guideline-making, particularly for preventive services in cancer (10). In order to assess the tradeoffs in health benefits and costs associated with a shift towards risk-stratified screening strategies, we used simulation modeling to evaluate novel adaptive cervical and colorectal cancer screening strategies that extend the screening interval based on one or more negative screening results. We aim to identify cost-effective approaches for cervical and colorectal cancer screening in the United States and highlight potential opportunities to improve cancer screening efficiency.

METHODS

Overview of Models

We utilized two previously developed microsimulation models, which have been the basis of several published cost-effectiveness analyses and screening evaluation studies in the

United States. Both models (Table 1) are part of the National Cancer Institute's Cancer Intervention and Surveillance Modeling Network (CISNET) consortium and have been used to inform USPSTF guidelines in the respective cancers.

For cervical cancer, we used an individual-based microsimulation model of HPV infections and cervical cancer natural history (11-14). The model simulates individual women from age 9 years until death through HPV, precancer (cervical intraepithelial neoplasia grades 2 and 3 (CIN2, CIN3)) and invasive cancer health states. Preclinical cancers may progress in stage or may be detected by symptoms. Transitions are based on multiple factors, such as age, duration of infection or precancer, HPV genotype and history of HPV infection. Previous literature includes details on the model's parameter estimation and calibration, as well as the data sources used in model development (13-15). Screening scenarios were adapted to allow for changes in screening interval based on screen results.

For colorectal cancer, we used the Simulation Model of Colorectal Cancer (SimCRC). SimCRC (16-21) simulates the development and growth of adenomas, some of which may progress to preclinical colorectal cancer. As with the cervical cancer model, SimCRC contains a screening component that simulates the detection of adenomas and preclinical cancers, which may progress in stage or be detected by symptoms. It also simulates removal of adenomas by polypectomy. The model was adapted for this analysis to allow screening intervals to be extended based on prior test results.

Table 1.1. Cervical and Colorectal model details and characteristics

Model	Cervical Model	Colorectal Model (SimCRC)
Approach	Microsimulation	Microsimulation
Time interval	Discrete (monthly)	Time-to-Event
Data Sources for Parameter Estimation/Calibration/ Validation	SEER, US Census, PROSPR, KPNC, NMHPVPR, CVT	SEER, US Census, NHANES, NHS, HPFS, NHIS
Natural History	Healthy, HPV (by genotype), CIN2, CIN3, Pre-clinical cancer (by stage), Clinical cancer (by stage)	Healthy, Low/Medium/High-risk Adenoma, Pre-clinical cancer (by stage), Clinical cancer (by stage)
Screening characteristics	Cytology† Sensitivity: 70 Specificity: 91 HPV Test (to detect HPV infection) ‡ Sensitivity: 91 Specificity: 93	FIT Sensitivity: - for adenomas 5mm or less: 5 - for adenomas 6-9mm: 10.1 - for adenomas 10mm or more: 22 - for cancer: 70 Specificity: 95
Outputs	Costs Cancer incidence Stage Distribution Test Results Mortality rate Quality-adjusted life expectancy	Costs Cancer incidence Stage Distribution Test Results Mortality rate Quality-adjusted life expectancy

† Cytology sensitivity and specificity values represent probabilities of ASCUS or worse/better given presence/absence of CIN 2 or worse health status.
‡ HPV DNA testing is assumed to be 100% sensitive and 100% specific in detecting the presence/absence of high-risk HPV types (pooled or by genotype). Under this assumption, the model generates an implied clinical sensitivity for detecting CIN 2 or worse of 91% and specificity of 93%.
Note:
* Test performance is assumed to be independent for multiple tests conditional on presence or absence of disease
SEER = Surveillance, Epidemiology, and End Results
PROSPR = Population-based Research to Optimize the Screening Process
KPNC = Kaiser Permanente Northern California
NMHPVPR = New Mexico HPV Pap Registry
CVT = Costa Rica Vaccine Trial
NHANES = National Health and Nutrition Examination Survey
NHS = Nurses' Health Study
HPFS = Health Professionals' Follow–up Study
NHIS = National Health Interview Survey

Screening Strategies

We evaluated cervical cancer screening strategies that were modifications of both

cytology and HPV co-testing guideline recommendations, allowing for extension of screening

intervals based on negative results (Table 1.1). For cytology-only screening, we extended

intervals to 5, 8, or 10 years after one or two consecutive negative screens. For co-testing, we

maintained a constant 3-year cytology testing interval between ages 21-29 years but evaluated

extending the co-testing interval from 5 years to 8, 10, 12, 15, or 20 years after one or two

consecutive negative co-tests beginning at age 30 years. We also included a strategy that

involved reducing the original co-testing interval to 3 years in order to mitigate some of the

health benefit tradeoff from extending screening. In total, we evaluated 31 cervical cancer

strategies, including no screening and both current guideline options (Table 1.2).

For colorectal cancer screening, we extended annual FIT screening to screening every 2

or 3 years following a range of 1 to 15 consecutive negative screens. Including the no-screen

scenario, we modeled 32 different colorectal screening strategies (Table 1.2).

Table 1.2. Description of Adaptive Screening Strategies, Cervical and Colorectal Cancer

Screening Type	Initial interval	Number of negatives needed to extend interval	New interval	Total strategies evaluated (including guidelines)
Cervical Cancer				
Cytology	Every 3 years	1, 2	5, 8, 10 years	7
Cytology/HPV co-test	Every 5 years (or 3 years)	1, 2	5, 8, 10, 12, 15, 20 years	23
Colorectal Cancer				
FIT	Annual	1-15	2, or 3 years	31

Analysis

Both models were used to evaluate the similar outcomes. Cancer cases, cancer deaths, and number of screening tests were estimated from the model to assess clinical and screening impact. In addition, health benefits and costs were estimated to evaluate cost-effectiveness. Health benefits were defined as quality-adjusted life years (QALYs) accrued over a lifetime horizon beginning at the age of screening initiation (21 years in the cervical cancer model and 50 years in the colorectal cancer model). Costs included the lifetime costs associated with cancer screening, diagnosis, and treatment of both precancer and invasive cancer. In adopting a societal perspective and keeping with recommended guidelines for cost-effectiveness analysis, we discounted both health benefits and costs annually at 3% and used a willingness-to-pay threshold of $100,000 per quality-adjusted life year (QALY) gained (22, 23). Consistent with assumptions in previous model-based analyses used to inform USPSTF recommendations, we assumed perfect adherence to screening.

To determine the optimal screening strategy for each cancer, we conducted an incremental cost-effectiveness analysis including all interventions. We first eliminated strategies that were more costly and less effective than a singular alternative strategy (strongly dominated) or a linear combination of two alternative strategies (weakly dominated). The remaining strategies created an efficiency frontier and allowed for the calculation of incremental cost-effectiveness ratios (ICERs) defined as the difference in cost divided by the difference in effectiveness between a strategy and the adjacent less costly strategy on the frontier. We compared these ICERs and considered the strategy with the highest ICER below a commonly-used cost-effectiveness threshold of $100,000 per QALY gained as the cost-effective cancer

screening strategy (23). We accounted for uncertainty in model parameters by conducting

sensitivity analyses on each model in areas where uncertainty had been identified from prior

analyses. For the cervical cancer model, all scenarios were run using 50 different good-fitting

natural history parameter sets. For colorectal cancer, we varied costs of screening, cancer care,

and treatment.

RESULTS

Non-dominated strategies

There were 12 efficient, or "non-dominated", strategies on the cervical cancer efficiency

frontier (Figure 1.1A), including the no-screening strategy and the guideline-based strategy

which involves switching to co-testing every 5 years at age 30. Non-dominated strategies include

all strategies for which no other strategy exists that is both cheaper and provides more QALYs.

Therefore, moving along the efficiency frontier, there exists a tradeoff between cost and benefit

(QALYs). The efficiency frontier for all non-dominated colorectal cancer screening strategies

(Figure 1.1B) contains 6 strategies, including the guideline-based FIT strategy.

Figure 1.1A/1B. Efficiency Frontier of Cervical and Colorectal Cancer Screening Strategies

Figure 1A: Cervical Cancer Efficiency Frontier

Figure 1B: Colorectal Cancer Efficiency Frontier
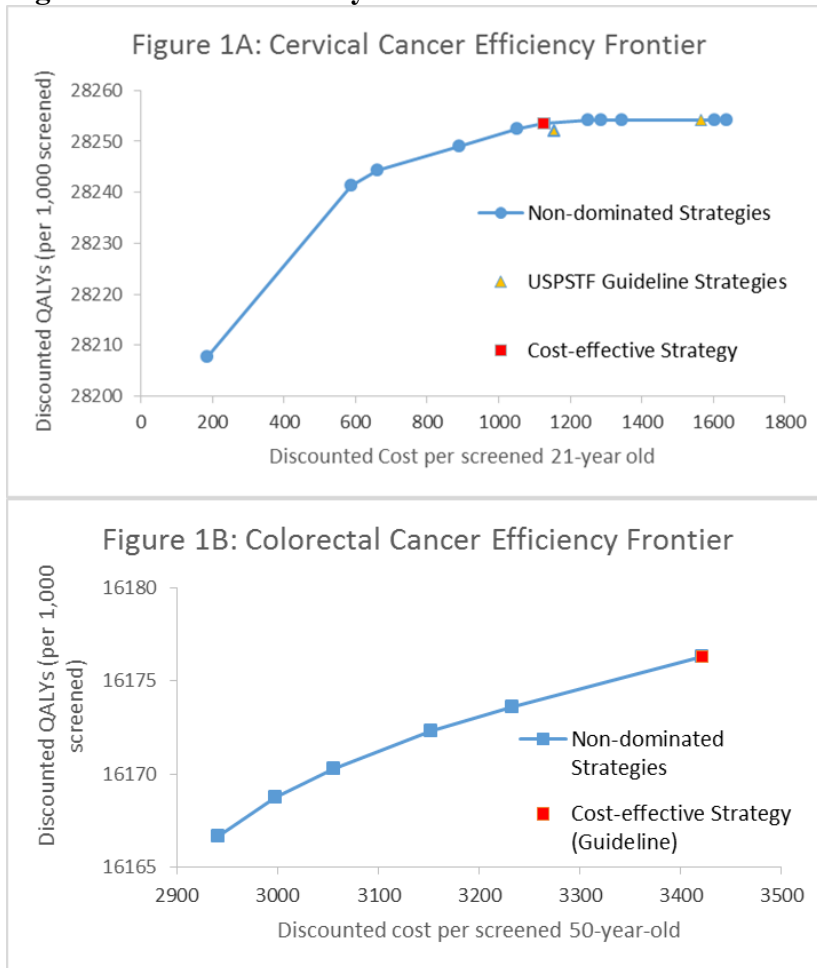
Cervical Cancer: Primary outcomes

Table 1.3 provides the primary outcomes of cancer cases, cancer deaths, screens, and diagnostic colposcopies across non-dominated adaptive strategies compared to guideline-based screening. The most effective strategy in minimizing cancer cases and deaths was to begin with a 3-year co-testing interval and extend it to 5 years after 1 negative screen. This strategy resulted in the same number of cancer deaths (33 per 100,000 screened) with one additional cancer case averted compared to guideline-based 5-year co-testing. Moving from guidelines to this strategy increased lifetime cytology screens by 41,000 (3.5%), lifetime HPV screens by 40,000 (4.7%) and lifetime diagnostic colposcopies by 4,000 (2.4%). A strategy which extended 5-year co-

testing to 8 years after 1 negative screen also resulted in 33 cancer deaths per 100,000 screened, but increased cancer cases to 99 per 100,000 screened compared to 97 per 100,000 screened in 5-year guidelines-based co-testing. However, compared to guidelines, this strategy reduced the number of cytology screens by 200,000 (18%), HPV screens by 200,000 (24.8%), and diagnostic colposcopies by 27,000 (16.3%).

Compared to 3-year cytology-only guidelines, which resulted in 223 cancer cases and 88 cancer deaths per 100,000 women an adaptive strategy extending 5-year co-testing to 15 years after 1 negative screen reduced cancer cases by 32.7% (73 cases per 100,000 screened) and cancer deaths by 37.8% (31 deaths per 100,000 screened). Additionally, the adaptive strategy reduced cytology screens by 642,000 (47.4%), while HPV screens increased by 258,000 (353%), and colposcopies increased by 37,000 (52%). Generally, the number of cancers cases and cancer deaths increased as the screening interval was extended further away from guidelines-based screening after one negative screen.

Table 1.3. Primary outcomes across selected non-dominated cervical screening strategies

Strategy*	Cancer Cases (per 100,000 screened)	Cancer Deaths (per 100,000 screened)	Lifetime Cytology Screens (per 100,000 screened)	Lifetime HPV Screens (per 100,000 screened)	Lifetime Colposcopies (per 100,000 screened)
Cervical Cancer					
Co-test – 3y – 5y – 2 neg	96	33	1,188,000	882,000	174,000
Co-test – 3y – 5y – 1 neg	96	33	1,155,000	848,000	170,000
Guideline 2 (Co-test 5y)	97	33	1,114,000	808,000	166,000
Co-test – 5y – 8y – 1 neg	99	33	914,000	608,000	139,000
Co-test – 5y – 10y – 2 neg	105	36	858,000	552,000	130,000
Co-test – 5y – 10y – 1 neg	110	40	821,000	515,000	125,000
Co-test – 5y – 15y – 1 neg	150	51	712,000	406,000	108,000
Guideline 1 (Cytology 3y)	223	82	1,354,000	73,000	71,000
Co-test – 5y – 20y – 1 neg	218	88	637,000	331,000	95,000
Cytology – 3y – 5y – 1 neg	328	123	973,000	53,000	52,000
Cytology- 3y – 8y – 1 neg	483	185	660,000	36,000	37,000
Cytology – 3y – 10y – 1 neg	589	226	554,000	30,000	31,000
No Screening	1,650	818	0	0	0

* Strategies are presented in the form (Screen Type – Original Interval in years – New Interval in years – Number of consecutive negative screens criterion required to change interval)

Colorectal Cancer: Primary outcomes

Adaptive colorectal cancer strategies yielded increasing cancer cases and deaths as the criteria for extending FIT screening from annual to biennial screening became more lenient (Table 1.4). Guideline annual FIT screening was the most effective in minimizing cancer cases and cancer deaths with 2,117 cases and 430 deaths, while lifetime FIT screens totaled 1,630,000 per 100,000 screened. Colonoscopies done for follow-up, surveillance, and symptomatic work-up totaled 179,000 per 100,000 screened under guideline annual screening. The closest adaptive strategy to guidelines was an adaptive strategy which extended the interval to 2 years after 14 negative screens. This strategy yielded 361 (2,478 vs 2,117) more cancer cases and 109 (539 vs 430) more cancer deaths per 100,000 individuals (increases of 17.1% and 25.3%, respectively).

However, lifetime FIT screens decreased by 16% (1,630,000 vs 1,370,000) and colonoscopies

decreased by 11.2% (179,000 vs 159,000) moving from guidelines to adaptive screening.

Table 1.4. Primary outcomes across selected non-dominated colorectal screening strategies

Strategy*	Cancer Cases (per 100,000 screened)	Cancer Deaths (per 100,000 screened)	Lifetime Screens (per 100,000 screened)	Lifetime Colonoscopies (per 100,000 screened)
Colorectal Cancer				
Annual FIT (Guideline)	2,117	430	1,630,000	179,000
FIT – 1y – 2y – 14 negs	2,478	539	1,370,000	159,000
FIT – 1y – 2y – 10 negs	2,612	568	1,270,000	151,000
FIT – 1y – 2y – 6 negs	2,772	603	1,160,000	142,000
FIT – 1y – 2y – 4 negs	2,870	625	1,100,000	137,000
FIT – 1y – 2y – 2 negs	2,990	646	1,030,000	131,000
No screening	6,951	2,774	0	6,951

Abbreviation: FIT, fecal immunochemical test
* Strategies are presented in the form (Screen Type – Original Interval in years – New Interval in years – Number of consecutive negative screens criterion required to change interval)

Cervical Cancer CEA

Table 1.5 gives costs, health benefits (QALYs), and incremental cost-effectiveness ratios

(ICERs), for the 12 non-dominated cervical screening strategies. At a willingness-to-pay

threshold of $100,000 per QALY gained, the strategy that extends the 5-year co-testing interval

to 15 years after a single negative screen (including both normal cytology and HPV-negative

results) at age 30 was cost-effective with an ICER of $62,700 This strategy strongly dominated

the cytology-only guidelines, providing greater health benefits at a lower cost. Compared to co-

testing guidelines, the cost-effective adaptive strategy resulted in a cost-savings of $439,000 per

1,000 women screened ($1,565,000 - $1,126,000), a savings of over 28%, with a decrease in

health benefits of 0.68 QALYs per 1,000 women screened, a reduction of 0.002% (Table 5).

Table 1.5. Cost-effectiveness of Cervical Cancer Screening Strategies

Strategy*	Cost (per 1,000 people screened)	QALYs (per 1,000 people screened)	Incremental Cost-Effectiveness Ratio (ICER), $
No Screening	184,000	28,208.66	-
Cytology – 21 – 3y – 10y – 1 neg	587,000	28,241.33	12,000
Cytology – 21 – 3y – 8y – 1 neg	662,000	28,244.28	25,400
Cytology – 21 – 3y – 5y – 1 neg	888,000	28,249.07	47,200
Co-test – 5y – 20y – 1 neg	1,051,000	28,252.35	49,700
Co-test – 5y – 15y – 1 neg	**1,126,000**	**28,253.54**	**62,700**
Co-test – 5y – 10y – 1 neg	1,250,000	28,254.08	227,900
Co-test – 5y – 10y – 2 neg	1,286,000	28,254.13	712,500
Co-test – 5y – 8y – 1 neg	1,345,000	28,254.17	1,356,000
Co-test – 5y (Guideline)	1,565,000	28,254.22	5,023,000
Co-test – 3y – 5y – 1 neg	1,605,000	28,254.22	7,243,300
Co-test – 3y – 5y – 2 neg	1,638,000	28,254.23	9,704,900

Abbreviation: QALY, quality-adjusted life year
* Strategies are presented in the form (Screen Type – Original Interval in years – New Interval in years – Number of consecutive negative screens criterion required to change interval)
** Assuming a willingness-to-pay of $100,000/QALY gained
*** Cost-effective strategy in bold

Colorectal Cancer CEA

Guideline-based annual FIT screening was the cost-effective (ICER of $70,300 screening strategy at the $100,000 threshold, providing 2 additional QALYs per 1,000 people screened compared to the next-best strategy at an additional cost of $188,000 (Table 1.6). At a lower willingness-to-pay threshold, such as $50,000 per QALY, the strategy to extend FIT screening from annual to bi-annual after 10 consecutive negative screens was cost-effective. This strategy yielded 4 less QALYs per 1,000 screened (0.02% reduction) compared to annual guideline screening at a cost-savings of $269,000 (7.9%) per 1,000 screened.

Table 1.6. Cost-effectiveness of Colorectal Cancer Screening Strategies

Strategy*	Cost (per 1,000 people screened)	QALYs (per 1,000 people screened)	Incremental Cost-Effectiveness Ratio (ICER)
FIT – 1y – 2y – 2 negs	2,941,000	16,167	-
FIT – 1y – 2y – 4 negs	2,998,000	16,169	$27,700
FIT – 1y – 2y – 6 negs	3,056,000	16,170	$37,600
FIT – 1y – 2y – 10 negs	3,152,000	16,172	$47,300
FIT – 1y – 2y – 14 negs	3,233,000	16,174	$63,400
Annual FIT (Guideline)	**3,421,000**	**16,176**	**$70,300**

Abbreviation: QALY, quality-adjusted life year; FIT, fecal immunochemical test
* Strategies are presented in the form (Screen Type – Original Interval in years – New Interval in years – Number of consecutive negative screens criterion required to change interval)
** Assuming a willingness-to-pay of $100,000/QALY gained
*** Cost-effective strategy in bold

Sensitivity Analysis

We conducted sensitivity analyses on uncertain model input parameters in both models. For cervical cancer, we ran all scenarios under 50 different good-fitting calibrated model input parameter sets as a form of probabilistic sensitivity analysis. Across all parameter sets, we found that the ranking ordering of strategies and the cost-effective strategy to be the same as in the base case analysis.

For colorectal cancer, we varied costs associated with screening and annual costs of cancer care. We increased costs by up to 50% for the diagnosis stage of patients diagnosed with stage 3 and 4 cancers, as well as total costs for those who die of cancer beginning at any stage. For all increases in costs, the ranking of strategies and the cost-effective strategy remained the same as in the base case. We also increased the cost of the FIT test, which had a base case cost of $39.48. We found that results were robust for increases in the FIT cost up to approximately

$61 (a 55% increase in cost), at which point the cost-effective strategy was an adaptive strategy in which annual FIT was changed to bi-annual after 14 negative screens (ICER $94,000 per QALY). Secondly, we ran the scenarios with increased FIT sensitivity for adenomas and colorectal cancer. For 10% and 20% increases in FIT sensitivity, we found no change in optimal strategy. However, for a 30% increase, an adaptive strategy to switch from annual to bi-annual after 10 negative screens was cost-effective (ICER $90,000 per QALY).

DISCUSSION

The USPSTF has relied on evidence from decision models in forming recent cancer screening recommendations (10). As data have become available on guidelines in practice, there exists an opportunity to further investigate the potential for efficiency gains through new risk-based strategies that had not been considered before. A major advantage of decision models is their ability to allow researchers to consider hypothetical scenarios and project the downstream costs and benefits once they are implemented. In this study, we evaluated screening protocols based on a novel approach that adapts screening interval length based on evidence provided by previous screening results. Furthermore, we show that in the case of cervical cancer, this adaptive approach would be cost-effective in the United States.

Risk-stratified screening strategies can improve cancer screening efficiency by identifying individuals at particularly low risk of developing cancer based on one or more early negative screens. By screening this population less often, we can significantly reduce screening costs and other potential harms associated with screening programs, with minimal effect on health outcomes. Two factors played a major role in the cost-effectiveness of adaptive screening

17

strategies. In cervical cancer screening, the high negative predictive value of the cytology/HPV co-test ensures that the loss in health benefits by lengthening the screening interval would be small. In colorectal cancer, on the other hand, the low sensitivity of the FIT makes it inefficient to lengthen the screening interval even after numerous consecutive negative tests. The second factor is test cost. In colorectal cancer screening, the cost of FIT is extremely low and, even when administered annually, was found to be a cost-effective way of mitigating the risk of delaying screening after a false negative.

The main drawback of the interval-extending screening strategies modeled in this study is the loss in health benefit due to decreased screening. However, Table 3 shows that the strategy which extends the co-testing interval to 8 years after 1 negative screen gives us nearly identical health benefits to current 5-year co-testing guidelines at a large cost-savings, with a health loss of only 0.05 QALYs per 1,000 women at a cost-savings of $220,000 per 1,000 women. Invested in other health interventions, these savings could be spent in other areas of health care to yield even more health improvements without any additional costs. For example, within cervical cancer prevention efforts, there may be an opportunity to re-invest these savings to more aggressively follow-up women with outcomes that classify them as higher risk or recruit women who are never screened or screened infrequently.

There are multiple limitations to this study in terms of modeling the proposed screening strategies. We assumed perfect adherence to screening for all scenarios. Though much evidence exists that screening adherence is not perfect (24, 25), we made this assumption following the base-case analyses used by the USPSTF in their recommendations. We did, however, conduct

sensitivity analyses on screening adherence to show that base-case results still hold under a change in our assumptions. An aspect to adherence that we did not model is the possibility that changing the screening interval may harm patient adherence. While there is evidence that longer screening intervals may not harm screening adherence (26, 27), it is possible that changing the interval for women at different times may impact adherence. Furthermore, in considering the quality-adjustment of life years during the screening program, our models do not account for utility decrements associated with the screening procedures themselves and false-positive results. Including these utility values would likely benefit interval-extending strategies that screen less often.

Other limitations involve aspects of cancer natural history that are unknown or unobservable and cannot be modeled precisely, which may impact the results of this analysis. For example, in cervical cancer, uncertainty around HPV re-infection in older women may impact the possibility of having additional infections which may be missed by extending screening intervals. In colorectal cancer, research indicates that cancer is increasing in younger populations and may have distinct characteristics from typical cancers, which are still unknown and difficult to model (28, 29). Furthermore, there is evidence of an alternative pathway to colorectal cancer through serrated polyps which could account for 10-20% of colorectal cancers (30). While these cancers are distinct from those caused by typical adenomas, there is still limited information regarding their natural history or how to address them in clinical practice (31). As a result, it is not clear how reducing the number of FIT tests, which do not detect these lesions well, may impact the effectiveness of screening in reducing these cancers (32). It is

possible false-positives FIT results may result in colonoscopies that improve the chance of detecting cancers that develop from a serrated polyp.

Given that risk-stratified screening strategies have the potential to improve cancer screening efficiency, an important next step will be to assess and address the challenges in adopting these strategies in clinical practice. One difficulty is that physicians are likely to be wary of how these strategies may adversely impact individual patients. If neither the cost of screening nor the savings from adopting an adaptive strategy are felt by the decision-maker, the cost-savings may not be enough to justify a change in behavior. However, it should be noted that if the cost-effective adaptive strategy for cervical cancer had been the default recommendation, switching to current guidelines-based screening would come at a cost of over $600,000 per QALY, well over commonly-cited cost-effectiveness thresholds in the United States. A second difficulty in implementation is that changing to an adaptive strategy creates additional protocol complexity and may result in low adherence from physicians, as well as patients who may have trouble keeping track of when they are supposed to complete their screening. However, increased use of electronic medical records (EMRs) may aid in future tracking of patient screening history, adjustments to screening schedules, and patient notifications which may mitigate some of the logistical issues.

In conclusion, a risk-based adaptive screening approach can improve efficiency in screening by accounting for patient heterogeneity in the screening population identified by prior screening results. In cervical cancer, we identified adaptive strategies that produced nearly equal benefit in terms of QALYs compared to more intensive non-adaptive guideline strategies at a

significant cost-savings. In colorectal cancer, we confirmed that the existing guidelines were cost-effective and identified features (test characteristics and costs) that reduce the attractiveness of an interval-extending approach. Considering adaptive, personalized screening strategies in future guideline development for cancer screening may help to reduce screening costs and fund other cost-effective interventions without placing a burden on the health care system.

CHAPTER 2

Cost-effectiveness Analysis of Updated ASCCP Guidelines for the Management of Women with
Abnormal Cervical Cancer Screening Results

ABSTRACT

BACKGROUND

Recently published 2019 guidelines by the American Society for Colposcopy and Cervical Pathology (ASCCP) changed the management of abnormal cervical cancer screening results from a result-based approach to a risk-based approach. We estimated the cost-effectiveness and changes in resource utilization of moving management of women screened for cervical cancer from 2012 to 2019 guidelines.

METHODS

We utilized a previously published model of cervical cancer costs and outcomes to estimate and compare the number of screens, colposcopies, treatments, cancer cases, and cancer deaths for the 2012 and 2019 ASCCP guidelines. We explored these guidelines under the scenarios of observed screening adherence and perfect screening adherence moving from 3-year cytology to both 3-year and 5-year cytology after age 30. In addition, we estimated lifetime costs and life years to determine the cost-effectiveness of moving to the updated risk-based guidelines.

RESULTS

Under observed screening compliance and perfect screening moving from 3-year cytology to 3-year co-testing at age 30, 2019 guidelines dominated 2012 guidelines by producing slightly greater life years at a cost-savings. For 3-year cytology screening moving to 5-year co-testing at age 30, 2019 guidelines were not cost-effective compared to 2012 guidelines at a willingness-to-pay threshold of $100,000 per life year (ICER $180,700). Across all scenarios, 2019 guidelines reduced the number of colposcopies and cancer deaths.

CONCLUSION

The 2019 ASCCP risk-based guidelines are likely to be a cost-effective and potentially cost-saving option compared to 2012 guidelines. "Equal management of equal risks" allows decision-makers to efficiently manage patients without having to work through complex algorithms that may reduce physician and patient adherence.

BACKGROUND

Over the past 20 years, cervical cancer screening guidelines have evolved, from annual cytology-based (pap smear) screening in 2003 and biennial cytology-based screening in 2009 to triennial cytology-based screening in 2012 (1, 33, 34). In addition, human papillomavirus (HPV) testing was included in the 2012 US Preventive Services Task Force (USPSTF) guidelines as part of "co-testing," which involves both cytology and HPV testing every 5 years in women over the age of 30; in the most recent guidelines in 2018, the screening options were expanded to include 5-year primary HPV testing alone after age 30 years as well (35). As the number of tests and, consequently, the number of possible test results has increased, the guidelines for management of women with unique screening results have also become more complex as decision-makers determine the best action for a particular combination of results. One simplifying principle is that of "equal management of equal risks" in which the strategy for follow-up of a particular woman depends on the risk that her past screen history implies (36, 37).

The American Society for Colposcopy and Cervical Pathology (ASCCP) is a professional organization consisting of physicians, researchers, and other stakeholders representing numerous societies, health organizations, and federal agencies. In 2002, the ASCCP published their first consensus guidelines on evidence-based management of abnormal cervical cancer screening results (38). As follow-up data on women who undergo screening have become more available, studies have been published detailing downstream risk of cervical disease and cancer for combinations of test results (3, 6, 39). Furthermore, organizations such as the USPSTF have established updated guidelines on screening interval and modality (1, 40). As a result, the

ASCCP updated their guidelines in 2006 and 2012, and recently published their 2019 guidelines (41-43). The goal of the newly revised update is to use the evidence base to increase the accuracy of risk estimates and actions for different test results and to simplify protocols that may have previously been too complex for both providers and patients (44-46).

Simulation modeling is a useful tool to assess the impact of future policy changes. This chapter aims to modify and utilize an existing model of cervical cancer to project changes in outcomes and estimate the cost-effectiveness of the new ASCCP guidelines for management of women with abnormal screening results. While the actions recommended by the 2012 guidelines were based on the specific test results, the new guidelines are risk-based and will allow decision-makers to more easily consider applying the "equal management of equal risks" principle. The results of this analysis will shed light on the efficiency of this new strategy and the potential impact it may have on health care resources and outcomes.

METHODS

Model

We used a previously developed model of cervical cancer that has been used extensively to evaluate the impact and cost-effectiveness of cervical cancer prevention strategies in the United States, including to inform recent USPSTF guidelines (11-14). Details of this model are included in Chapter 1 of this dissertation and in previous publications (13-15) but are briefly summarized here. States of the model are depicted in Figure 2.1. Simulated women transition through states on the basis of factors such as age, duration of infection or precancer, or

genotype/history of HPV infection. These states include a normal, healthy cevix which can be infected with HPV, as well as two precancerous states, CIN2 (cervical epithelial neoplasis grade 2) and CIN3 (CIN grade 3). From these states, progression to cervical cancer is possible. Screening strategies and management protocols for women with abnormal screening results can be modified within the model, which then produces relevant outputs of cost, precancer incidence, cancer incidence, life years, and test counts.

Figure 2.1. Markov structure of cervical cancer model
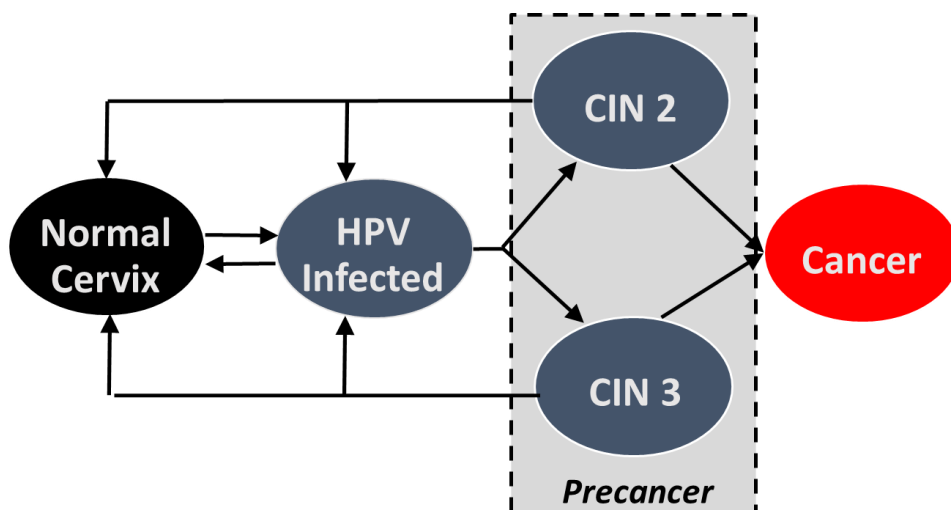
Note: Death state is not shown here, but mortality is possible from any of the states depicted

Validation/Calibration

Data from Kaiser Permanente Northern California (KPNC) used to inform the 2019 ASCCP guidelines provided estimates of cumulative CIN3+ risk (the probability that a woman develops CIN grade 3 or cervical cancer) for distinct sets of index screening results, including multiple rounds of screening (3, 6, 39). Empirical risk data from KPNC are provided in Table 2.1, with action thresholds determined by the ASCCP provided in the Supplemental Material. In

the model, we tracked 5-year CIN3+ risk for several combinations of co-testing results from

primary screening to ensure that risks inferred by testing results in the model appropriately fit

KPNC observed data. Person-level outputs were manipulated to calculate 5-year CIN 3+

estimates. To account for discrepancies in model fit, we adjusted input parameters that affected

the distribution of test results until visual fit near or within the 95% confidence intervals of the

data was achieved.

Table 2.1. 5-year CIN3+ risk observed among women ages 21-65 years from Kaiser Permanente Northern California, stratified by preceding co-test result.

Co-test Result (Cytology/HPV)	N	Proportion of co-test results (%)	CIN3+ 5 year risk (%)	95% CI
NILM/Neg	1388153	89.76	0.12	(0.12, 0.13)
ASCUS/Neg	25331	1.64	0.40	(0.31, 0.48)
LSIL/Neg	3300	0.21	1.96	(1.40, 2.52)
HSIL/Neg	183	0.01	27.37	(20.27, 34.46)
NILM/Pos	63541	4.11	4.80	(4.58, 5.03)
ASCUS/Pos	30506	1.97	7.27	(6.87, 7.67)
LSIL/Pos	23659	1.53	6.94	(6.49, 7.40)
HSIL/Pos	3980	0.26	53.16	(51.18, 55.13)

Abbreviation: CIN, cervical intraepithelial neoplasia; NILM, negative for intraepithelial lesion or malignancy; ASCUS, atypical squamous cells of undetermined significance; LSIL, low-grade squamous intraepithelial lesion; HSIL, high-grade squamous intraepithelial lesion; Neg, negative HPV test; Pos, positive HPV test

Model Scenario

New ASCCP guidelines contained several changes from the previous 2012 update,

requiring development of a unique screening and management strategy in the model. Table 2.2

summarizes key differences in management following first round of co-test screening results.

The three primary differences are (1) for a "NILM/Negative" screen, an extension from a 3-year

routine follow-up to 5-years in updated guidelines, (2) for a "LSIL/Negative" result, a switch

from colposcopy to 1-year follow-up, and (3) for a "HSIL/Positive" result, a decision to refer

women to immediate precancer treatment.

Table 2.2. Summary of 2012 and updated 2019 guidelines for management of women with normal and abnormal cervical cancer screening results

1st Test Result (Cyto/HPV)	NILM/ Neg	ASCUS/ Neg	LSIL/ Neg	HSIL/ Neg	NILM/ Pos	ASCUS/ Pos	LSIL/ Pos	HSIL/ Pos
New Guidelines	5-year	3-year	1-year	Colpo	1-year	Colpo	Colpo	Treatment
2012 Guidelines	3-year	3-year	Colpo	Colpo	1-year HPV Test	Colpo	Colpo	Colpo

Abbreviation: CIN, cervical intraepithelial neoplasia; NILM, negative for intraepithelial lesion or malignancy; ASCUS, atypical squamous cells of undetermined significance; LSIL, low-grade squamous intraepithelial lesion; HSIL, high-grade squamous intraepithelial lesion; Colpo, Colposcopy; Neg, negative HPV test; Pos, positive HPV test

In addition to results from the first round of screening, we utilized KPNC data on a

second round of screening, as well as test results conditional on preceding colposcopy result

and/or treatment (Table 2.3). For example, Table 2 indicates that the interval between screens be

reduced to 1 year for an LSIL/HPV-Negative result. However, as seen in Table 2.3, if an

LSIL/HPV-Negative result follows a NILM/HPV-negative result, 5-year CIN3+ risk indicates

reducing screening to 3-year intervals instead of an annual interval. In the model, once a third

screening result is observed, the first result is no longer considered, and the action is based upon

the two most recent screens. Similar to actions from a single round of screening, key changes in

the new guidelines were primarily around added scenarios for sending women straight to treatment after a "HSIL/Positive" result, additional scenarios with 1-year follow-up rather than colposcopy, and reduced intensity for results following a "NILM/Negative" or "ASCUS/Negative" in the first round of screening.

Table 2.3. Two-test combination actions as advised by updated ASCCP guidelines

1st Test Result	2nd Test Result							
	NILM/Neg	ASCUS/Neg	LSIL/Neg	HSIL/Neg	NILM/Pos	ASCUS/Pos	LSIL/Pos	HSIL/Pos
NILM/Neg	5-year	3-year	3-year	Colpo	1-year	1-year	1-year	Colpo
ASCUS/Neg	3-year	1-year	1-year	Colpo	1-year	1-year	1-year	Colpo
LSIL/Neg	3-year	1-year	1-year	Colpo	1-year	1-year	1-year	Colpo
ASC-H/Neg	1-year	1-year	1-year	Colpo	1-year	1-year	1-year	Colpo
HSIL/Neg (followed by <CIN2)	Colpo + Cotest	1-year	1-year	Colpo	1-year	1-year	1-year	Colpo
HSIL/Neg (followed by CIN2)	1-year x3	1-year x3	1-year x3	Colpo	Colpo	Colpo	Colpo	Treatment
HSIL/Neg (followed by CIN3)	1-year x3	1-year x3	1-year x3	Colpo	Colpo	Colpo	Colpo	Treatment
Neg/Pos	1-year	1-year	1-year	Colpo	1-year	1-year	1-year	Colpo
ASCUS/Pos (followed by <CIN2)	3-year	1-year	1-year	Colpo	1-year	1-year	1-year	Colpo
ASCUS/Pos (followed by CIN2)	1-year x3	1-year x3	1-year x3	Colpo	Colpo	Colpo	Colpo	Treatment
ASCUS/Pos (followed by CIN3)	1-year x3	1-year x3	1-year x3	Colpo	Colpo	Colpo	Colpo	Treatment
LSIL/Pos (followed by <CIN2)	3-year	1-year	1-year	Colpo	1-year	1-year	1-year	Colpo
LSIL/Pos (followed by CIN2)	1-year x3	1-year x3	1-year x3	Colpo	Colpo	Colpo	Colpo	Treatment
LSIL/Pos (followed by CIN3)	1-year x3	1-year x3	1-year x3	Colpo	Colpo	Colpo	Colpo	Treatment
HSIL/Pos (followed by <CIN2)	Colpo + Cotest	1-year	1-year	Colpo	Colpo	Colpo	Colpo	Treatment
HSIL/Pos (followed by CIN2)	1-year x3	1-year x3	1-year x3	Colpo	Colpo	Colpo	Colpo	Treatment
HSIL/Pos (followed by CIN3)	1-year x3	1-year	1-year	Colpo	Colpo	Colpo	Colpo	Treatment

Abbreviation: Colpo, Colposcopy; CIN, cervical intraepithelial neoplasia; NILM, negative for intraepithelial lesion or malignancy; ASCUS, atypical squamous cells of undetermined significance; LSIL, low-grade squamous intraepithelial lesion; HSIL, high-grade squamous intraepithelial lesion; Neg, negative HPV test; Pos, positive HPV test

Note: 1-year refers to annual screening; 1-year x3 implies annual screening until 3 consecutive NILM/Neg results are observed before return to 3-year screening

Analysis

There were two primary outcomes of this analysis. First, we conducted a cost-effectiveness analysis of the new guidelines compared to previous guidelines. We simulated 1,000,000 women under three different screening scenarios. We mimicked current screening practice as observed in the KPNC population with respect to screening interval, compliance to diagnostic colposcopy/biopsy, and compliance to precancerous treatment (47-49). Second, we conducted analyses assuming perfect adherence to screening guidelines in which women undergo 3-year cytology-based screening until age 30 followed by 3-year co-testing until age 65, as practiced in the KPNC system. Finally, we included scenario analysis assuming 3-year cytology-based screening until age 30 followed by 5-year co-testing until age 65, consistent with USPSTF guidelines. We determined the cost-effective strategy in each screening scenario using standard willingness-to-pay thresholds of both $50,000 and $100,000 per life-year gained (LYG). For both costs and health benefits, we applied an annual discount rate of 3%, as is recommended for cost-effectiveness studies in the United States (22, 23). Additionally, we compared resource utilization and outcomes across each scenario to determine the extent to which the new guidelines impact number of screens, colposcopies, treatments, and cancers.

RESULTS

Validation

Figure 2.2 shows observed and modeled 5-year CIN3+ risk, ordered by how often each outcome occurs. The top graph shows screening outcomes that are observed in more than 1% of screens, with the NILM/HPV-Negative result occurring in 89.8% of cases. The bottom graph

shows outcomes which occurred in less than 0.5% of screens. These include HSIL/HPV-Positive, LSIL/HPV-Negative, and HSIL/HPV-Negative. Model output slightly overestimates risk for women with ASCUS/HPV-positive results (2.0% of total results), and slightly underestimates LSIL/HPV-positive (1.0%) and HSIL/HPV-negative (0.01%) results. Overall trends across testing results for model risk follow observed data.

Figure 2.2. Comparison of 5-year CIN3+ risk between observed data and model estimates, stratified by first co-test result
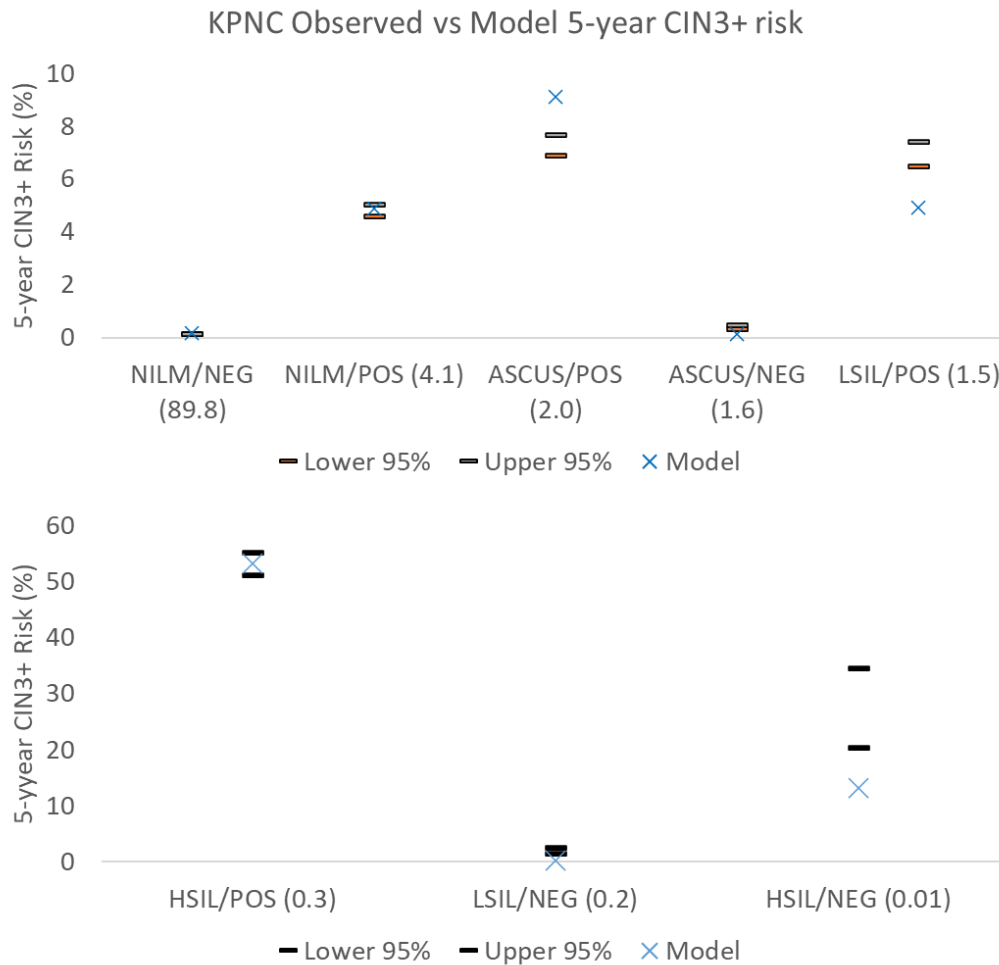
Abbreviation: CIN, cervical intraepithelial neoplasia; NILM, negative for intraepithelial lesion or malignancy; ASCUS, atypical squamous cells of undetermined significance; LSIL, low-grade squamous intraepithelial lesion; HSIL, high-grade squamous intraepithelial lesion; Neg, negative HPV test; Pos, positive HPV test
Note: Parentheses indicate percentage of total population

Resource utilization and cancer outcomes

Table 2.4 displays utilization of screens, colposcopies and treatments, as well as total cancers detected and cancer deaths in each screening scenario. Under observed screening, the revised guideline, which was the cost-effective strategy, resulted in 2.15 million fewer screens (including both HPV and cytology tests), 410,000 fewer colposcopies, 202 fewer cancers detected, and 114 fewer cancer deaths per million women. However, the number of treatments under the new guidelines increased by nearly 3,000 per million women screened. The scenario under which 3-year cytology switches to 3-year co-testing at age 30 followed the same trend, saving 650,000 screens, 390,000 colposcopies, 209 cancers, and 96 cancer deaths. However, treatments in this scenario increased by 3,150. Finally, under perfect screening switching from 3-year cytology to 5-year co-testing at age 30, screens increased by 2.93 million, while colposcopies reduced by 200,000 and treatments reduced by nearly 6,000 when shifting from the previous guidelines to the new updated guidelines. While the number of detected cancers under the new guidelines increased by 153, the number of cancer deaths reduced by 127.

Table 2.4. Model estimates for resource utilization and cancer outcomes between 2019 (New) and 2012 (Old) ASCCP guidelines per million women

Observed screening	Screens (Cytology + HPV, millions)	Colposcopies (millions)	Treatments (thousands)	Cancers Detected	Cancer Deaths
New Guidelines	31.44	1.12	163.09	1,410	423
Old Guidelines	33.59	1.53	160.15	1,612	537
Difference (New – Old)	**-2.15**	**-0.41**	**2.94**	**-202**	**-114**
Perfect 3y/3y screening					
New Guidelines	30.14	1.10	159.36	1,435	435
Old Guidelines	30.79	1.49	156.21	1,644	531
Difference (New – Old)	**-0.65**	**-0.39**	**3.15**	**-209**	**-96**
Perfect 3y/5y screening					
New Guidelines	25.47	1.00	150.98	1,486	433
Old Guidelines	22.54	1.20	156.87	1,333	560
Difference (New – Old)	**2.93**	**-0.20**	**-5.89**	**153**	**-127**

Cost-effectiveness Analysis

Table 2.5 gives results of the cost-effectiveness analysis between new and old guidelines under multiple screening scenarios. Assuming screening adherence as observed in KPNC, the new 2019 ASCCP risk-based guidelines dominate 2012 guidelines, providing additional life years at a cost savings of nearly $134 per person. Under the perfect screening scenario in which women are screened with cytology-based screening every 3 years until age 30 before switching to 3-year co-testing, the 2012 guidelines are still dominated by 2019 guidelines with an additional cost of around $87 at a loss in life years. However, assuming perfect screening in which women switch to 5-year co-testing at age 30, 2019 guidelines are no longer the cos-effective management strategy at the $100,000 per life year threshold compared to 2012

guidelines. New guidelines cost an additional $72 per person, while adding .0004 life years per person for an ICER of $180,700 per life year compared to old guideline management.

Table 2.5. Cost-effectiveness analysis comparing new and old ASCCP guidelines under observed screening, perfect screening switching from 3-year cytology to 3-year co-testing at age 30, and perfect screening switching from 3-year cytology to 5-year co-testing at age 30

Observed screening practice	Cost per person ($, discounted)	Difference	Life Years per person (years, discounted)	Difference	ICER ($/LY)
New Guidelines	1900		28.2534		
Old Guidelines	2030	133.75	28.2532	-0.0002	Dominated
Perfect 3y cytology/3y co-testing screening					
New Guidelines	1780		28.2527		
Old Guidelines	1870	86.56	28.2525	-0.0001	Dominated
Perfect 3y cytology/5y co-testing screening					
Old Guidelines	1500		28.2523		
New Guidelines	1570	72.29	28.2526	0.0004	180,700

Note: Within each group, scenarios are ordered by cost

Sensitivity Analysis

We conducted cost-effectiveness analysis across all three scenarios for 50 different good-fitting natural history parameter sets. Across all 50 sets, 2019 guidelines were cost-effective for observed KPNC and perfect KPNC (3-year cytology followed by 3-year co-testing) screening scenarios. For observed KPNC, 2019 guidelines dominated 2012 guidelines under 49 of 50 parameter sets, with 1 parameter set resulting in an ICER of over $1,000,000 per life year. Under perfect KPNC screening, 2019 guidelines dominated under all 50 parameter sets.

However, under perfect USPSTF screening (3-year cytology followed by 5-year co-testing), 2019 guidelines were no longer cost-effective compared to 2012 guidelines, with ICERs ranging from $100,500 to $279,000 per life year across all parameter sets. Difference in resource utilization were also robust across all parameter sets.

DISCUSSION

Based on the risks calculated from the observed KPNC data and the thresholds for various actions determined by the ASCCP, it is evident that three main differences exist between previous and new guidelines. First, the new guidelines take a more conservative screening approach for those who test negative and may be at a significantly reduced risk for future cervical disease. Second, the new guidelines also take a more conservative approach to some intermediate results to follow-up after one year instead of opting for colposcopy, particularly when the abnormal test immediately follows a negative co-test. While colposcopies are useful in identifying both cancer and precancer, they are invasive and a potentially unnecessary use of resources for patients under a certain risk threshold. Finally, the new guidelines take a more aggressive action against the worst screening outcome, HSIL/HPV-positive, to treat immediately.

The results of this analysis indicate that the risk-based approach introduced by the ASCCP improves the efficiency of cervical cancer management. Under both scenarios of observed screening and perfect screening switching from 3-year cytology to 3-year co-testing at age 30, new guidelines are cost-effective at any willingness-to-pay threshold, dominating 2012

guidelines. Furthermore, the life years produced between the two strategies are nearly identical, but the newer guidelines are dramatically cost saving. This finding suggests that any health decrement from reducing the number of screens and colposcopies is balanced by the additional treatments resulting from the more aggressive approach. On the other hand, under the scenario in which women switch from 3-year cytology to 5-year co-testing at age 30, 2019 guidelines are no longer cost-effective, likely due to the decreased number of routine screens compared to the increased number of 1-year surveillance screens in the new guidelines. One benefit of the more aggressive approach to treatment is to benefit those who are not as compliant with screening, so it may not be surprising to see that this benefit is reduced when screening compliance is assumed to be perfect.

There are a number of limitations to be considered in this analysis. We validated risk estimates for results from an initial round of screening but did not consider test results conditional on prior colposcopy results and/or treatment. While some observed data are available on these outcomes, for some combinations of results, the number of observations across each outcome may not be enough to confidently determine if the model estimates are accurate. In addition, the KPNC data represent a population of well-screened women who may not be representative of the national population to generalize the results of this analysis. Finally, we did not report outcomes considering quality of life associated with screening procedures or cancer as data on quality adjustment for increased cervical cancer testing and results are variable and difficult to integrate accurately in our model. Given that the number of cancer deaths associated with the 2012 guidelines is higher in the 3-year cytology switching to 5-year co-

testing case, a severe decrement to quality of life for late stage cancer may make newer guidelines more appealing.

The latest ASCCP guidelines are an important step in the management of women who screen positive for cervical disease. "Equal management of equal risks" allows decision-makers to efficiently manage patients based on risk as determined by testing history without having to work through complex algorithms that may reduce physician and patient adherence. The results produced in this analysis support that the new ASCCP guidelines are likely to be a cost-effective and potentially cost-saving option, compared to prior guidelines that relied on algorithm-based management. Moving forward, it will be important to present these new guidelines in a simple and easy-to-implement way that can facilitate greater compliance and benefit both the health care system and patient outcomes.

CHAPTER 3

Imputation and Decision Modeling to Improve Diagnosis and Management of Patients at Risk for Post-Transplant Diabetes Mellitus

ABSTRACT

BACKGROUND

Post-transplantation diabetes mellitus (PTDM) is an important complication of solid organ transplantation with limited representation in the literature. More studies are needed to explore solutions for better diagnosis and management of the population at risk for this disease. In this study, we determine the extent to which PTDM goes undiagnosed over the course of 1 year following transplantation, analyze missed or later-diagnosed cases of PTDM due to poor Hemoglobin A1c (HbA1c) and fasting blood glucose (FBG) collection, and estimate the impact a screening intervention to better collect glucose metrics may have on long-term outcomes.

METHODS

This was a retrospective study utilizing three datasets from a single center. The kidney transplantation dataset consisted of 407 patients who underwent transplant between 1999 and 2006. The liver dataset consisted of 346 patients who underwent transplant between 2007 and 2012. Finally, the heart dataset contained 152 patients who underwent transplant between 2010 and 2015. Retrospective analysis was supplemented with an imputation procedure to account for missing data and project outcome under perfect information. In addition, the data was used to inform a simulation model used to estimate life expectancy and cost-effectiveness of a hypothetical intervention.

RESULTS

The estimate of PTDM incidence increased from 27% to 31% in kidney transplantation patients, 31% to 40% in liver transplantation patients, and 45% to 67% in heart transplantation patients,

when HbA1c and FBG were assumed to be collected perfectly at all time points. Simulated life expectancy for kidney transplantation patients was 18.97 under perfect screening at a cost savings compared to 18.84 years in current practice. For liver transplantation patients, perfect data collection increased life expectancy by 0.10 years (9.60 vs 9.50, ICER $37,000). In heart transplantation patients, the benefit of perfect screening was 0.44 years (11.06 vs 10.62, ICER $30,800).

CONCLUSION

Improved collection of HbA1c and FBG is a cost-effective method for catching many additional cases of PTDM within the first year alone. Additional research into both improved glucometric monitoring as well as effective strategies for mitigating PTDM risk will become increasingly important to improve health in this population

BACKGROUND

Post-transplantation diabetes mellitus (PTDM) is one of the potential unwelcome outcomes following solid organ transplantation affecting patients without a prior history of a diabetes mellitus (DM) diagnosis. This newly-diagnosed DM can lead to decreased graft function, increased risk of cardiovascular disease, and lower survival, resulting in increased downstream healthcare costs (50, 51). Reported estimates of PTDM incidence vary widely: 4-25% in kidney transplantations, 2.5-25% in liver transplants, and 4-40% in hearts transplantations (52-56). Over 35,000 solid organ transplantations occurred in the United States in 2018, marking the sixth consecutive year in which the number of transplantations increased from the previous year (57). The extent to which PTDM burdens the healthcare system and solutions to improve diagnosis, management, and specialized treatment has not been well-studied.

A 2014 international expert panel released updated recommendations for the management of transplant patients at-risk for PTDM (58). Among their recommendations was the expansion of screening tests for PTDM using hemoglobin A1c (HbA1c) and blood glucose monitoring. There is currently no standardized protocol for regular PTDM screening in the post-transplantation period and the potential impact of this recommendation is unknown. Studies of PTDM have utilized collection of HbA1c and blood glucose as determinants for diagnosis of PTDM. However, collection of HbA1c and fasting blood glucose (FBG) is inconsistent and unavailable for analysis in many patients across multiple follow-up time points (59-61). In some cases, patients may have had their DM undiagnosed due to lack of data collection or diagnosed

months after an early screen would have been indicated. As a result, it is likely that PTDM incidence is largely underestimated, and that the time course of PTDM development may be different from what is understood in the current literature (61). In addition, management specifically geared towards diagnosed PTDM remains vague. Most guidelines for decision-making follow broad type-2 DM protocols and specify importance of graft survival over concern for PTDM risk (58). However, the transplant patient population differs in many ways from the broad diabetes population (e.g., in the ability to modify exercise and diet in the immediate post-transplantation period) and thus standard type 2 DM management methods may need modification for this population (62).

In the face of uncertainty and lack of evidence to inform decision-making, quantitative methods such as data imputation and disease modeling can help to make use of available data and estimate potential for improved outcomes. Data imputation can allow observations from inadequate or incomplete data to be utilized in estimating missing data points and extending outcomes to a longer time horizon than currently available. Imputation in both clinical trial studies and observational studies, including those involving transplantation data and imputation of glycemic indicators, has become more commonplace (61, 63-67). In addition, disease and decision models can simulate outcomes for a patient population and utilize data on intervention effectiveness from different sources to project outcomes, thereby providing a decision maker with additional evidence to inform decisions. Decision modeling has been previously used, particularly with respect to DM, to estimate life expectancy, costs, and disease progression (68-71).

In this study, we explore a screening strategy for improvement of PTDM diagnosis and potential treatment. First, we use data imputation to estimate the impact of a screening intervention that perfectly monitors HbA1c and FBG in the follow-up period post transplantation. We analyze data on all major solid organ transplantations (kidney, liver, and heart) to determine (a) the extent to which PTDM goes undiagnosed over the course of 1 year following transplantation, and (b) cases of PTDM that occur later than would have been predicted if HbA1c and FBG had been screened at all follow-ups. Secondly, we develop a disease model of PTDM outcomes to project life expectancy and estimate the impact a screening intervention to better collect glucose metrics may have on long-term outcomes. This study represents the first modeling analysis of an intervention tailored to the transplantation population at risk of PTDM and should help decision makers in determining how effective we might expect such an intervention to be in improving health outcomes.

METHODS

Study Population

We utilized 3 previously published data sets on kidney, liver, and heart transplantation patients (59-61, 72, 73). The datasets were compiled through a de-identified chart review with IRB approval. The kidney transplantation dataset consisted of 407 patients who underwent transplant between 1999 and 2006. The liver dataset consisted of 346 patients who underwent transplant between 2007 and 2012. Finally, the heart dataset contained 152 patients who underwent transplant between 2010 and 2015. The heart dataset included information on demographics and medical history, as well as HbA1c, FBG, and some lab values at 1-, 2-, 3-, 4-, 6-, 8- and 12-months post-transplant. The kidney and liver patients were followed at 1-, 4-, and

12-months post-transplant and data were collected on the same variables as the heart patients, as well as hypoglycemic medication administered to patients at each time point post-transplant. Immunosuppression protocols differed by organ and are described in our previous studies on these patient populations (59, 60).

PTDM definition and Imputation

Patients were classified with PTDM using two standard definitions. First, we defined PTDM patients as those who met at least one of the criteria of FBG > 126 mg/dL or HbA1c $\geq$ 6.5% at one of the follow-up time points. Because collection of HbA1c and blood glucose was not consistent across all patients and follow-ups, we also classified a patient as having PTDM if they were being treated with insulin at a given follow-up visit. While kidney and liver datasets contained 1-, 4-, and 12-month follow-ups, the heart data set contained many additional follow-ups and inconsistent data collection across them. Therefore, as in previous studies utilizing this datasets, we combined follow-up months 1-3, 4-6, and 8-12 into 3 follow-up "periods" for analysis (60). To proxy perfect collection of HbA1c and FBG, we used Multiple Imputation by Chained Equations (MICE) to estimate missing values for these variables at each follow-up point. MICE has previously been used in the literature to replace missing data in the kidney dataset (61). Because the heart data contained less variables than the kidney and liver datasets, the regression equation used to impute only contained available HbA1c and FBG at each time point and the use of insulin in the final 24 hours before discharge as the predictors for the missed HbA1c and FBG values. For the kidney and liver datasets, the available HbA1c and FBG values, as well as the use of insulin at each of the follow-up time points, was used.

Simulation model overview and structure

We developed a microsimulation model that simulated patients who had undergone transplantation in the United States to estimate aggregated costs and life expectancy post-transplant. Patients were simulated on an annual cycle and could transition between one of six post-transplantation states: (1) a healthy state representing normal glycemic control; (2) an "Untreated PTDM" diabetes state defined as fasting blood glucose (FBG) greater than 125 mg/dL or HbA1c greater than 6.4%; (3) a "Treated PTDM" diabetes state representing someone who has been diagnosed with PTDM, but whose glycemic indicators are under control through treatment; (4) a graft rejection state representing patients who had a graft rejection and were either re-transplanted or, in the case of some kidney transplantation patients, assumed to be on dialysis for the remainder of their lifetime; (5) A re-transplanted state representing patients who undergo an additional transplant after graft failure; and (6) death. The Markov chain model representation of these states is provided in Figure 3.1.

Figure 3.1. A Markov Chain Model of PTDM
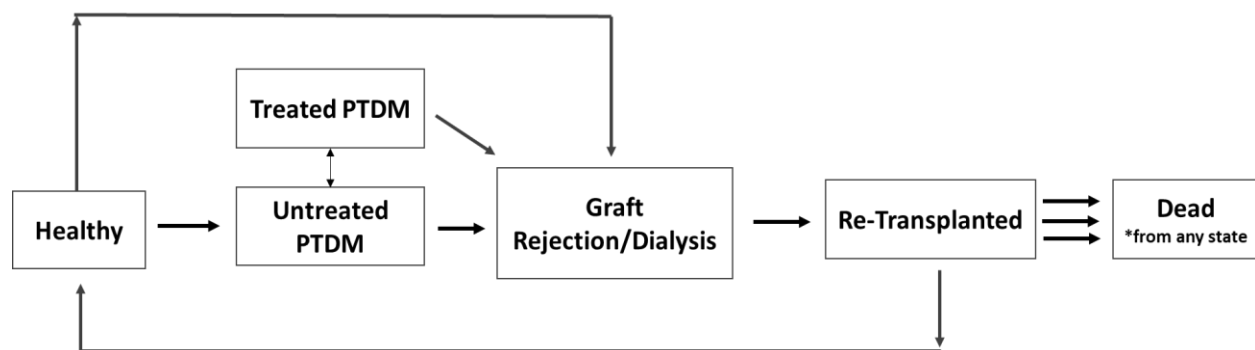
Note: Graft rejection state is possible for all 3 organs; Dialysis only represents kidney transplant recipients; Dead state can be reached from any of the other 5 states.

We used observed data on PTDM status in the immediate post-transplant setting to determine initial state probabilities, particularly the proportion of patients who enter the model in the "Healthy" or "Untreated PTDM" state. The difference in this proportion between the observed and perfect screening scenarios represents the benefit of early hyperglycemia detection by HbA1c and FBG monitoring. Transition probabilities were estimated by calculating probabilities of changing health states from the observed data. To account for uncertainty, we calculated a 95% confidence interval assuming a binomial distribution for each point estimate and applied these probabilities in sensitivity analyses.

As long-term outcomes were not available in the data, mortality rates or relative mortality risks for each health state were obtained from the published literature (50, 74-79). A summary of model parameters utilized is given in Table 3.1. For the screening analysis, we ran 50 Monte Carlo simulation cohorts of 100,000 individuals aged 40, 50, and 60 years-old entering the model and aggregated life years accrued and costs under both the current and perfect screening scenarios. We conducted a cost-effectiveness analysis utilizing a commonly accepted willingness-to-pay threshold of $100,000 per life year (23). We conducted a sensitivity analysis on the mortality and diabetes risk reduction due to earlier PTDM detection, a variable on which we currently have no data to inform the model. The model was built and analyses conducted using TreeAge Pro 2019 software.

Table 3.1. PTDM Model parameters estimated from data and literature

Parameter	Kidney (Annual rates)	Liver	Heart	Source
Mortality rate (Healthy post-transplant)	.00527 (age 20-39) .0193 (age 40-59) .0505 (age 60+)	0.035 (18-49) 0.050 (50-64) 0.066 (65+)	.051 (<35, >65) .042 (35-64)	Morales et al.(80)
Relative mortality risk (PTDM vs Healthy)	1.87 (1.60, 2.18)	1.29 (1.21, 1.36)	1.3 (1.1, 1.5)	Kasiske et al; Wong et al; Deo et al(50, 81, 82)
Relative mortality risk (Dialysis vs Healthy)	.0026 (age 20-39) .0118 (age 40-59) .0193 (age 60+)	N/A	N/A	Kaballo et al.(79)
Transition Probabilities	**Annual Probability**			
Healthy to PTDM	0.033 (.009, .082)		0.018 (.0005,.096)	Observed Mayo Data
Healthy to Rejection	0.028 (.006, .079)	N/A	N/A	Observed Mayo Data
PTDM to Healthy	0.51 (.358, .663)		0.659 (.500, .795)	Observed Mayo Data
PTDM to Rejection	0.045 (.001, .228)	N/A	N/A	Observed Mayo Data
Rejection to Re-retransplant	0.1	N/A	N/A	Observed Mayo Data
Costs				
Post-transplant (Annual)	$16,844			
Organ rejection (one-time)	$70,581			
Re-transplantation (one-time)	$106,373			
Additional cost of DM (annual)	8,500 (Age 18-40) 3,400 (Age >41)			
FBG screen	$5.25			
HbA1c screeen	$61			

Imputation Analysis

We first report means and confidence intervals for HbA1c, FBG, and insulin use at all time points/periods in both the observed and imputed (perfect collection) datasets for all three organs. We restricted the datasets to just those patients who did not have hyperglycemia or a diagnosed DM prior to transplant. Insulin use was not available for the heart dataset and therefore not included. Logistic regression models were built for each organ's dataset to determine the strength of association between available patient characteristics and risk factors with probability of missing HbA1c or FBG values at the 1-month follow-up. Each regression model was adjusted for age, sex, race, BMI, pre-transplant HbA1c, pre-transplant FBG, mean inpatient glucose, and transplant year. Finally, we compared observed PTDM diagnosis at each follow-up with PTDM diagnosis assuming perfect screening and data collection in the imputed dataset and estimated life expectancy under observed and perfect screening conditions.

RESULTS

Kidney

Table 3.2 gives descriptive statistics for kidney transplantation patients on HbA1c, FBG, and insulin use at 1-, 4-, and 12-months post-transplant for both the observed and imputed data sets. Means for the imputed dataset were identical to observed data. Mean HbA1c across the three follow-ups rose, while FBG decreased over time. The percentage of patients on insulin increased from 11.9% at 1 month to 13.9% at 4 months before dropping to 9.9% at 12 months. The number of missing HbA1c values decreased over time from 106 at 1 month to 53 at 12 months. FBG collection was perfect at 1 month, but 14 and 29 patients had missing values at 4 and 12 months, respectively.

Table 3.2. Observed and imputed HbA1c, FBG, and insulin use at 1-, 4-, and 12-months post-transplant for kidney transplantation patients who did not have pre-transplant DM

Variable	n	Observed Data Missing values	Mean	95% CI	n	Imputed Perfect Collection Missing values	Mean	95% CI
1-month HbA1c	197	106	5.6	(5.5, 5.6)	303	0	5.6	(5.5, 5.6)
4-month HbA1c	203	100	5.6	(5.6, 5.7)	303	0	5.6	(5.6, 5.7)
12-month HbA1c	250	53	5.8	(5.7, 5.9)	303	0	5.8	(5.7, 5.9)
1-month FBG	303	0	112	(107, 117)	303	0	112	(107, 117)
4-month FBG	289	14	103	(101, 106)	303	0	103	(101, 106)
12-month FBG	274	29	103	(101, 106)	303	0	103	(101, 105)
On insulin at 1 month	303	0	11.9%	(8.2, 15.5)	303	0	11.9%	(8.2, 15.5)
On insulin at 4 months	303	0	13.9%	(9.9, 17.8)	303	0	13.9%	(9.9, 17.8)
On insulin at 12 months	303	0	9.9%	(6.5, 13.3)	303	0	9.9%	(6.5, 13.3)

Liver

The 1-, 4-, and 12-month observed and imputed values for liver transplantation patients on HbA1c, FBG, and insulin use are found in Table 3.3. Similar to the data collected from kidney transplant patients, the number of missing HbA1c values drop from 204 at 1 month to 76 at 4 months, before increasing slightly to 95 at 12 months. In addition, missing FBG values increased from 11 to 14 to 29 going from 1 month to 4 months to 12 months, respectively. Insulin use was much higher in liver patients compared to kidney, and also decreased over time, moving from 51.5% at 1 month to 21.5% at 12 months.

Table 3.3. Observed and imputed HbA1c, FBG, and insulin use at 1-, 4-, and 12-months post-transplant for liver transplantation patients who did not have pre-transplant DM

Variable	n	Observed Data Missing values	Mean	95% CI	n	Imputed Perfect Collection Missing values	Mean	95% CI
1-month HbA1c	33	204	5.5	(5.3, 5.6)	237	0	5.4	(5.2, 5.5)
4-month HbA1c	161	76	5.5	(5.4, 5.6)	237	0	5.5	(5.4, 5.5)
12-month HbA1c	142	95	5.7	(5.6, 5.9)	237	0	5.6	(5.5, 5.8)
1-month FBG	226	11	100	(97, 103)	237	0	100	(97, 102)
4-month FBG	223	14	105	(102, 108)	237	0	105	(102, 108)
12-month FBG	208	29	111	(106, 115)	237	0	111	(107, 115)
On insulin at 1 month	237	0	51.1%	(44.6, 57.5)	237	0	51.1%	(44.6, 57.5)
On insulin at 4 months	237	0	33.8%	(27.7, 39.8)	237	0	33.8%	(27.7, 39.8)
On insulin at 12 months	237	0	21.5%	(16.2, 26.8)	237	0	21.5%	(16.2, 26.8)

Heart

Table 3.4 gives descriptive statistics for heart transplantation patients on aggregated HbA1c and FBG across the 1-3 month, 4-6 month, and 8-12 month follow-up periods. FBG data was nearly complete across all 3 time periods in this cohort. However, 82 (81%) of HbA1c values were missing in the 1-3 month period. Missingness improved to just 73 (72%) by 4-6 months and 58 (57%) by 8-12 months. Mean HbA1c and FBG both decreased over time. Compared to the observed data set, mean imputed HbA1c values were lower in both the 1-3 month (6.6 vs 6.1) and 4-6 month (6.7 vs 5.8) periods.

Table 3.4. Observed and imputed HbA1c and FBG across 1-3month, 4-6month, and 8-12month follow-up periods post-transplant for heart transplantation patients who did not have pre-transplant DM

Variable		Observed Data				Imputed Perfect Collection		
	n	Missing values	Mean	95% CI	n	Missing values	Mean	95% CI
1-3 month HbA1c	19	82	6.6	(6.0, 7.2)	101	0	6.1	(5.9, 6.3)
4-6 month HbA1c	28	73	6.7	(6.1, 7.2)	101	0	5.8	(5.5, 6.1)
8-12 month HbA1c	43	58	5.6	(5.3, 5.8)	101	0	5.6	(5.4, 5.8)
1-3 month FBG	99	2	125	(118, 131)	101	0	125	(118, 131)
4-6 month FBG	98	3	107	(103, 111)	101	0	107	(103, 111)
8-12 month FBG	99	2	105	(101, 110)	101	0	105	(101, 110)

Long-term outcomes

Table 3.5 summarizes diagnosis of PTDM by each follow-up time point/period across all organs for both the observed and imputed datasets. Perfect collection of HbA1c and FBG resulted in 12 (15%) more PTDM cases in the kidney cohort, 21 (28%) more cases in the liver cohort, and 23 (51%) more cases in the heart cohort. In terms of time course of PTDM diagnosis, a majority (75%) of extra PTDM diagnoses in the kidney cohort occurred at the 12-month time period. In the liver cohort, however, 16/21 (76%) of extra cases were identified by 1 month, while 16/23 (70%) of extra post-heart transplant cases were identified in the 1-3 month time period. Overall, PTDM incidence using these cohorts increased from 27% to 31% in the kidney group, 31% to 40% in the liver group, and 45% to 68% in the heart group assuming imputation (using MICE) as a proxy for full screening and data collection. These findings suggest that previous studies in which imputation was not used to handle missing values might have underestimated PTDM incidence.

Table 3.5. PTDM diagnosis by follow-up time period across kidney, liver, and heart transplantation patients for both observed and imputed datasets

	Kidney				Liver				Heart*			
	Observed		Perfect Collection		Observed		Perfect Collection		Observed		Perfect Collection	
	n	DM	n	DM	n	DM	n	DM	n	DM	n	DM
1 month	303	63	303	64	228	18	237	34	99	39	101	55
4 months	233	11	239	14	210	32	203	26	59	5	46	9
12 months	207	8	225	16	163	24	177	35	55	1	37	4
TOTAL PTDM Diagnoses	**82**		**94**		**74**		**95**		**45**		**68**	
Estimated PTDM incidence	**27%**		**31%**		**31%**		**40%**		**45%**		**67%**	

Note: PTDM incidence may differ from previously published studies utilizing these datasets as patients were excluded who did not have available data or had defined hyperglycemia in the pre-transplant period (compared to previous studies which only excluded patients with a diagnosed DM)

* Heart dataset contains 1-3, 4-6, and 8-12 month time periods

Early Hyperglycemic Screening as a Potential Remedy

We now make use of simulation to estimate the impact of early screening as a potential remedy that can improve patient outcomes by allowing for previously undiagnosed or late-diagnosed cases to be treated earlier. Figure 3.2 represents simulated life expectancy outcomes using base case parameters and results of a sensitivity analysis on effectiveness of early hyperglycemia detection in bringing patients to a state of "Treated" PTDM for kidney, liver, and heart, respectively. For kidney transplantations (Figure 3.2A), assuming perfect identification and treatment of PTDM, life expectancy in the perfect screening case is approximately 18.97 years compared to 18.84 years in the observed case, a difference of 0.13 years, or 1.6 months. This gain in life years occurs at a cost-savings, making perfect screening the dominant strategy. We also conducted a threshold analysis to determine, how effective our ability to identify DM

patients and treat would need to be in order for perfect screening to remain cost-effective at a $100,000 per life year threshold. We find that, though the base case increase is 21.1% from 7.3% in the observed group, only 8.5% of patients need to begin in the "Treated" state in order for the screening scenario to be cost-effective.

In addition to base case parameters, we also conducted a sensitivity analysis in which data-derived parameter values were varied across their 95% confidence interval. Due to the small sample size available, we chose to vary these parameters using a uniform distribution to obtain a wide range of outcomes. We ran 50 cohorts of 100,000 patients and output life expectancy for current practice compared to perfect screening for each cohort. Across all 50 runs, the screening scenario resulted in longer life expectancy than current practice. For 40-year-olds, screening resulted in a life expectancy of 23.17 years (95% Credible interval (CI) 23.11-23.23) compared to 23.03 years in current practice (22.96-23.12). For 50-year-olds, life expectancy with screening was 19.04 years (18.96-19.10) compared to 18.89 years (18.78-18.94) in current practice. Finally, 60-year-olds had a screening life expectancy of 13.87 years (13.79-13.95) compared to 13.64 years (13.57-13.72) in current practice.

Figure 3.2. Life expectancy for post-transplant kidney, liver, and heart transplantation patients under observed and perfect screening scenarios for 50-year old patients
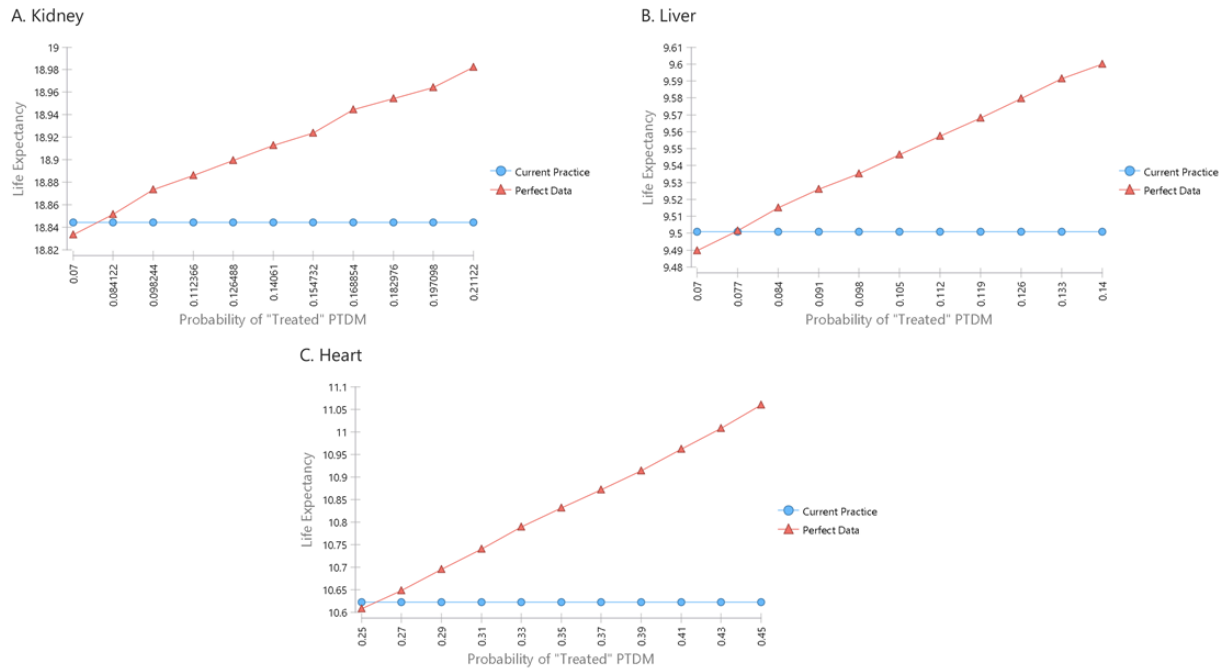

A. Kidney

B. Liver

C. Heart

Figure 3.2B displays impact of early PTDM detection and successful control for liver transplantation recipients. Perfect screening would result in a 0.10 year (9.60 vs 9.50 years) increase in life expectancy per person at a cost difference of $3,700 resulting in an incremental cost-effectiveness ratio (ICER) of $37,000 per life-year gained. Perfect identification and treatment of PTDM assumes 14% of patients being treated. However an increase from 7.6% in the current practice case to just 9% would make the perfect screening strategy cost-effective at the $100,000 threshold. Across 40-year-olds, the perfect screening scenario gave an average life expectancy of 10.74 (10.69 – 10.81) years compared to 10.64 (10.57 – 10.70) years in current practice. 50-year-olds gained also gained 0.1 years per person screened from 9.75 (9.66 – 9.83) with current practice to 9.85 years (9.78 – 9.95) with perfect screening. Finally, 60 year-olds

with perfect screening gained 0.11 years from 9.05 years (8.97 – 9.13) in current practice to 9.16

(9.10 – 9.23) with perfect screening.

Incidence estimates for heart transplantation data were impacted the most from

information gained through perfect screening, with an absolute increase of 22% in PTDM

incidence (Table 3.5). Much of this increase was seen in the first month, where the estimate of

immediate PTDM rose from 38% to 54%. Figure 3.2C shows the benefit obtained by the earlier

detection of these extra PTDM cases in the base-case scenario, with a life expectancy of 11.06

years in the perfect screening scenario compared to 10.62 in current practice at a cost of $13,500

(ICER $30,800 per life-year gained). Virtually any improvement in ability to identify and treat

patients with PTDM in the perfect screening scenario was cost-effective. Discounted life

expectancies for 40-year-old patients were 11.04 years (10.98 – 11.15) under current information

compared to 11.48 years (11.39 – 11.56) under perfect information. For 50-year-old patients, the

difference was 0.45 years with a life expectancy of 10.75 years (10.65 -10.82) versus 11.2 years

(11.10 – 11.28). Finally, 60-year-old patients lived on average 10.12 years (10.07 – 10.19) under

current practice compared to 10.59 years (10.51 – 10.65) under perfect screening.

Table 3.6. Logistic regression odds ratios for association between pre-transplant and inpatient patient characteristics and probability of missing variables

	Kidney		Liver		Heart	
	1-month HbA1c	1-month FBG	1-month HbA1c	1-month FBG	1-month HbA1c	1-month FBG
Age, per 5 years	.91 (.76, 1.08)	N/A	.94 (.73, 1.21)	1.17 (.78, 1.75)	1.06 (.84, 1.33)	N/A
BMI	1.04 (.95, 1.14)	N/A	1.02 (.95, 1.10)	.96 (.86, 1.07)	1.00 (.88, 1.12)	N/A
Race (White)	.80 (.30, 2.16)	N/A	.44 (.10, 2.03)	1.3 (.15, 11.4)	**4.45 (1.43, 13.86)***	N/A
Male (vs Female)	.83 (.33, 2.10)	N/A	.56 (.22, 1.38)	.36 (.11, 1.25)	.92 (.27, 3.10)	N/A
Pre-transplant HbA1c	1.58 (.28, 8.90)	N/A	.69 (.31, 1.53)	.47 (.13, 1.70)	N/A	N/A
Pre-transplant FBG	1.01 (.98, 1.05)	N/A	N/A	N/A	1.03 (.98, 1.08)	N/A
Mean inpatient glucose	.95 (.76, 1.20)	N/A	1.02 (.83, 1.27)	.97 (.68, 1.38)	.74 (.42, 1.31)	N/A
Transplant year (compared to year after)	**5.90 (3.72, 9.47)***	N/A	**1.39 (1.10, 1.75)***	.92 (.64, 1.33)	1.10 (.78, 1.57)	N/A

Note: 95% confidence intervals given in parentheses; 1-month FBG values collected cully in Kidney and Heart cohorts; Pre-transplant FBG and HbA1c unavailable liver cohort
* denotes significant values (all p<0.01)

Predictors of missing values

Odds ratios for the logistic regressions that are used to determine associations between pre-transplant patient characteristics and the probability of missing data values are given in Table 3.6. In both the kidney and liver datasets, transplant year was a significant predictor of missing HbA1c at 1 month. For each additional year from 2006 to 1999, having a kidney transplant one year earlier was associated with a 590% (OR 5.90, 95% CI 3.72, 9.38, p<.01) increase in the odds of missing data collection of HbA1c at 1 month. Among liver patients, moving from 2012 to 2007, having transplant a year earlier was associated with a 139% (OR 1.39, 95% CI 1.10-1.75, p<.01) increase in odds of missing 1-month HbA1c data. In addition, white heart transplant patients had a 445% (OR 4.45, 95% CI 1.43-13.86, p<.01) increased odds of missing 1-month HA1c. No other factors were significant.

DISCUSSION

Our data from kidney, liver, and heart transplantation patients demonstrate differences in HbA1c and FBG collection across organs as well as inconsistency in collection within a particular organ cohort at each follow-up time point. Protocols do not currently exist for standardized collection of these data, but improved collection of these measures can benefit in multiple ways. First, improved screening leads to earlier detection and better outcomes. The results of this analysis, at the very least, support design of a study to compare current practice to wide scale HbA1c and FBG monitoring in transplantation patients to monitor potential development of PTDM and study of patient characteristics that may determine who benefits greatest for a more targeted intervention. On a larger scale, better collection of these data improves estimates of PTDM burden and understanding of the need to prioritize research in this area. The 2014 Consensus Guidelines on PTDM and other literature describing research in PTDM treatment and management have expressed a need for more prevention strategies and interventions to treat newly diagnosed patients (56, 58). This analysis contributes to this area by providing estimates for how improved collection of glycemic measures post-transplant may enhance our understanding, and consequently, our ability to best manage those at risk of PTDM.

Across all three organs, FBG collection was relatively high for all time points, ranging from 88% to 100% of patients. In contrast, HbA1c collection was very limited. Around two-thirds of patients had HbA1c collected at 1-month and 4-month in the kidney cohort, while collection in the liver cohort was only 14% at 1-month before rising to 68% at 4-months. Despite

combining heart transplant patient data into multi-follow up time periods, only 19% of patients'

HbA1c was collected in the 1-3 month period, rising to 28% in the 4-6 month period, and 43% in

the 8-12 month period.

The microsimulation model developed for this analysis allowed for an estimation of life

years aggregated over the lifetime following transplantation. This model allowed us to

determine an estimate for the potential effectiveness of a screening-based intervention to collect

HbA1c and FBG at all follow-up time points. Estimates for life expectancy gained on average in

the perfect screening base case ranged from 0.10 years (1.2 months) in liver patients to 0.44

years (5.3 months) in heart transplant patients. While some patients may not benefit at all from

this intervention, this value applied to the full population of transplantation patients may

represent a meaningful improvement in health. Furthermore, we show that the low cost of both

HbA1c and FBG screening make monitoring of these values at each follow-up very cost-

effective at a $100,000 per life year threshold.

There are important limitations to note around this study, particularly regarding the use of

an imputation as a proxy for data collection. The reliability and value of the MICE imputation is

dependent on the different variables and data points available to inform the regression. All three

data sets had many missing HbA1c data points. While Tables 3.1-3.3 indicate that the

imputation did not substantially affect means and standard deviations, the distribution of the

imputed values may differ from the underlying unobserved data. The regression conducted and

reported in Table 3.4 indicates that HbA1c collection was not missing in a systematic manner

related to another available variable. However, other unobserved or uncollected risk factors may

have played a role in explaining why some patients had HbA1c and FBG collected at a follow-up while others did not. Additionally, the heart dataset did not contain information on insulin use at each time point. Therefore, we do not know if patients who had PTDM early on and missing values later on were not actually on continued insulin therapy.

There are also a number of limitations with respect to the PTDM simulation model. First, the model itself makes a number of assumptions regarding the limited number of health states and the lack of intermediate levels of glucose tolerance (such as pre-diabetes) which may have impact on mortality and other co-morbidities. There may also be other risk factors or aspects of DM natural history which we are not including. While one of the strengths of decision models, parameters were utilized from a number of different sources where data were collected from different types of populations that may be inherently different. While this heterogeneity of data sources builds some uncertainty in model results, it still represents the best available information. In addition, data sources for validation of life expectancy outputs were difficult to gather as PTDM in this populatoin is under-represented in the literature and few studies exist detailing outcomes of a generalizable population. Therefore, the purpose of this model was simply to use differential mortality rates between those without DM, those with DM, and those on dialysis, as well as transition probabilities between these states (all of which are observed), to gain an understanding of the potential order of magnitude change in life expectancy using perfectly collected (imputed) data and cost-effectiveness for screening in this population.

There are two broad implications of this study. First, we find that perfect collection of HbA1c and FBG results in catching many additional cases of PTDM within the first year alone.

This is significant because, as seen in Table 3.5, it indicates that estimates of PTDM incidence in the published literature may be significantly underestimating its burden and time-course. Perfect monitoring and early detection of previously undiagnosed or late-diagnosed PTDM cases is likely to be a cost-effective intervention to improve life expectancy in this population.

Second, this simulation model represents the first model of health outcomes for the post-transplantation population with regards to PTDM, a disease that imposes a significant burden in the transplantation community. There are important extensions to this modeling work that should be applied, including evaluation of new diagnostic, surveillance, and treatment strategies to mitigate PTDM risk while considering any tradeoffs to risk of organ rejection or other comorbidities. Additional research into both improved glucometric monitoring as well as safe and effective strategies for mitigating PTDM risk will become increasingly important to understand the impact of potential interventions on our ability to detect and treat early DM cases and improve health in this patient population.

Supplemental Material

Chapter 2

Supplement Figure 2.1. ASCCP risk threshold and recommended action for management of women with abnormal test results

Cervical Cancer Risk	Action
<0.15% 5-year CIN3+ risk	Return in 5 years
0.15-0.54% 5-year CIN3+ risk	Return in 3 years
>0.55% 5-year CIN3+ risk	Return in 1 year
4-24% immediate CIN3+ risk	Colposcopy recommended
25-59% immediate CIN3+ risk	Expedited treatment or colposcopy acceptable
>60% immediate CIN3+ risk	Expedited treatment preferred

Supplemental Figure 2.2 Observed Kaiser Permanente Northern California (KPNC) screening adherence by interval length

Screening Interval	Proportion of population
Age 21-30 (Cytology only)	
1 year	0.16
2 years	0.362
3 years	0.334
4 years	0.091
5+ years	0.053
Age 30-65 (Co-testing)	
1 year	0.036
2 years	0.144
3 years	0.633
4 years	0.133
5+ years	0.054

Chapter 3

Supplemental Table 3.1. Comparison of characteristics between renal, liver, and heart transplantation patients without pre-transplant diabetes.

Characteristic	Kidney (n = 291)	95% CI	Liver (n =250)	95% CI	Heart (n = 109)	95% CI
Age, y	49 (15)	(47, 51)	54 (9)	(53, 55)	51 (12.8)	(49, 54)
Male, %	56	(51, 62)	67	(61, 73)	69%	(60, 78)
White, %	71	(66, 76)	88	(84, 92)	69%	(60, 78)
Pre-transplant BMI, kg/m^2	26.8 (5.6)	(26.1, 27.4)	28.2 (5.6)	(27.5, 28.9)	25.9 (4.5)	(25.1, 26.8)
Live donor (%)	68	(62, 73)	22	(17, 27)		
Pre-transplant HbA1c, %	5.4 (0.3)	(5.4, 5.5)	5.2 (0.5)	(5.1, 5.3)	5.75 (.5)	(5.64, 5.86)
Steroid Use at 1 month, %	47	(42, 53)	98	(97, 100)		
Insulin Use at 1 month, %	6	(3, 9)	36	(30, 42)		
Inpatient mean glucose post-transplant, mg/dL	132 (19)	(129, 134)	148 (17)	(145, 150)	134 (10)	(132, 136)
Mean glucose 24 hours before hospital discharge, mg/dL	136 (39)	(132, 141)	144 (28)	(140, 148)	134 (23)	(129, 139)
Developed PTDM by 1 year, %	19		30		38	

Reported values are mean (SD) or percentages
Note: Tables adapted from previously published literature (59, 60)

REFERENCES

1. Moyer VA, Force USPST**.** Screening for cervical cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2012;156(12):880-91, W312.

2. Force USPST, Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson KW, Epling JW, Jr., et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. JAMA. 2016;315(23):2564-75.

3. Katki HA, Kinney WK, Fetterman B, Lorey T, Poitras NE, Cheung L, et al. Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: a population-based study in routine clinical practice. Lancet Oncol. 2011;12(7):663-72.

4. Dijkstra MG, van Zummeren M, Rozendaal L, van Kemenade FJ, Helmerhorst TJ, Snijders PJ, et al. Safety of extending screening intervals beyond five years in cervical screening programmes with testing for high risk human papillomavirus: 14 year follow-up of population based randomised cohort in the Netherlands. BMJ. 2016;355:i4924.

5. Demarco M, Lorey TS, Fetterman B, Cheung LC, Guido RS, Wentzensen N, et al. Risks of CIN 2+, CIN 3+, and Cancer by Cytology and Human Papillomavirus Status: The Foundation of Risk-Based Cervical Screening Guidelines. J Low Genit Tract Dis. 2017;21(4):261-7.

6. Castle PE, Kinney WK, Xue X, Cheung LC, Gage JC, Zhao FH, et al. Effect of Several Negative Rounds of Human Papillomavirus and Cytology Co-testing on Safety Against Cervical Cancer: An Observational Cohort Study. Ann Intern Med. 2017.

7. Singh H, Turner D, Xue L, Targownik LE, Bernstein CN**.** Risk of developing colorectal cancer following a negative colonoscopy examination: evidence for a 10-year interval between colonoscopies. JAMA. 2006;295(20):2366-73.

8. Zauber A, Knudsen A, Rutter C, Naber S, Doria-Rose P, Pabiniak C, et al. Evaluating the Benefits and Harms of Colorectal Cancer Screening Strategies: A Collaborative Modeling Approach; 2015.

9.   Chesson HW, Ekwueme DU, Saraiya M, Watson M, Lowy DR, Markowitz LE. Estimates of the

      annual direct medical costs of the prevention and treatment of disease associated with human

      papillomavirus in the United States. Vaccine. 2012;30(42):6016-9.

10.   Owens DK, Whitlock EP, Henderson J, Pignone MP, Krist AH, Bibbins-Domingo K, et al. Use of

      Decision Models in the Development of Evidence-Based Clinical Preventive Services

      Recommendations: Methods of the U.S. Preventive Services Task Force. Ann Intern Med.

      2016;165(7):501-8.

11.   Kim JJ, Burger EA, Sy S, Campos NG. Optimal Cervical Cancer Screening in Women Vaccinated

      Against Human Papillomavirus. J Natl Cancer Inst. 2017;109(2).

12.   Kim JJ, Campos NG, Sy S, Burger EA, Cuzick J, Castle PE, et al. Inefficiencies and High-Value

      Improvements in U.S. Cervical Cancer Screening Practice: A Cost-Effectiveness Analysis. Ann

      Intern Med. 2015;163(8):589-97.

13.   Campos NG, Burger EA, Sy S, Sharma M, Schiffman M, Rodriguez AC, et al. An updated natural

      history model of cervical cancer: derivation of model parameters. Am J Epidemiol.

      2014;180(5):545-55.

14.   Kim JJ, Ortendahl J, Goldie SJ. Cost-effectiveness of human papillomavirus vaccination and

      cervical cancer screening in women older than 30 years in the United States. Ann Intern Med.

      2009;151(8):538-45.

15.   Kim JJ, Kuntz KM, Stout NK, Mahmud S, Villa LL, Franco EL, et al. Multiparameter calibration of a

      natural history model of cervical cancer. Am J Epidemiol. 2007;166(2):137-50.

16.   Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, Kuntz KM.

      Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S.

      Preventive Services Task Force. Ann Intern Med. 2008;149(9):659-69.

17. Knudsen AB, Hur C, Gazelle GS, Schrag D, McFarland EG, Kuntz KM**.** Rescreening of persons with a negative colonoscopy result: results from a microsimulation model. Ann Intern Med. 2012;157(9):611-20.

18. Lansdorp-Vogelaar I, Gulati R, Mariotto AB, Schechter CB, de Carvalho TM, Knudsen AB, et al. Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits. Ann Intern Med. 2014;161(2):104-12.

19. Knudsen AB, Zauber AG, Rutter CM, Naber SK, Doria-Rose VP, Pabiniak C, et al. Estimation of Benefits, Burden, and Harms of Colorectal Cancer Screening Strategies: Modeling Study for the US Preventive Services Task Force. JAMA. 2016;315(23):2595-609.

20. van der Steen A, Knudsen AB, van Hees F, Walter GP, Berger FG, Daguise VG, et al. Optimal colorectal cancer screening in states' low-income, uninsured populations-the case of South Carolina. Health Serv Res. 2015;50(3):768-89.

21. Goede SL, Kuntz KM, van Ballegooijen M, Knudsen AB, Lansdorp-Vogelaar I, Tangka FK, et al. Cost-Savings to Medicare From Pre-Medicare Colorectal Cancer Screening. Med Care. 2015;53(7):630-8.

22. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. JAMA. 2016;316(10):1093-103.

23. Neumann PJ, Cohen JT, Weinstein MC**.** Updating cost-effectiveness--the curious resilience of the $50,000-per-QALY threshold. N Engl J Med. 2014;371(9):796-7.

24. Cuzick J, Myers O, Hunt WC, Robertson M, Joste NE, Castle PE, et al. A population-based evaluation of cervical screening in the United States: 2008-2011. Cancer Epidemiol Biomarkers Prev. 2014;23(5):765-73.

25.   Eltoum IA, Roberson J**.** Impact of HPV testing, HPV vaccine development, and changing screening frequency on national Pap test volume: projections from the National Health Interview Survey (NHIS). Cancer. 2007;111(1):34-40.

26.   Rendle KA, Schiffman M, Cheung LC, Kinney WK, Fetterman B, Poitras NE, et al. Adherence patterns to extended cervical screening intervals in women undergoing human papillomavirus (HPV) and cytology cotesting. Prev Med. 2017.

27.   Ogilvie GS, Smith LW, van Niekerk D, Khurshed F, Pedersen HN, Taylor D, et al. Correlates of women's intentions to be screened for human papillomavirus for cervical cancer screening with an extended interval. BMC Public Health. 2016;16:213.

28.   Ballester V, Rashtak S, Boardman L**.** Clinical and molecular features of young-onset colorectal cancer. World J Gastroenterol. 2016;22(5):1736-44.

29.   Willauer AN, Liu Y, Pereira AAL, Lam M, Morris JS, Raghav KPS, et al. Clinical and molecular characterization of early-onset colorectal cancer. Cancer. 2019;125(12):2002-10.

30.   Patai AV, Molnar B, Tulassay Z, Sipos F**.** Serrated pathway: alternative route to colorectal cancer. World J Gastroenterol. 2013;19(5):607-15.

31.   Kim SY, Kim TI**.** Serrated neoplasia pathway as an alternative route of colorectal cancer carcinogenesis. Intest Res. 2018;16(3):358-65.

32.   Greuter MJ, Xu XM, Lew JB, Dekker E, Kuipers EJ, Canfell K, et al. Modeling the Adenoma and Serrated pathway to Colorectal CAncer (ASCCA). Risk Anal. 2014;34(5):889-910.

33.   American College of O, Gynecologists**.** ACOG practice bulletin. Cervical Cytology screening. Number 45, August 2003. Int J Gynaecol Obstet. 2003;83(2):237-47.

34.   Warren JB, Gullett H, King VJ**.** Cervical cancer screening and updated Pap guidelines. Prim Care. 2009;36(1):131-49, ix.

35.   Force USPST, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Screening for Cervical

        Cancer: US Preventive Services Task Force Recommendation Statement. JAMA.

        2018;320(7):674-86.

36.   Castle PE, Sideri M, Jeronimo J, Solomon D, Schiffman M. Risk assessment to guide the

        prevention of cervical cancer. J Low Genit Tract Dis. 2008;12(1):1-7.

37.   Castle PE, Katki HA. Screening: A risk-based framework to decide who benefits from screening.

        Nat Rev Clin Oncol. 2016;13(9):531-2.

38.   Wright TC, Jr., Cox JT, Massad LS, Twiggs LB, Wilkinson EJ, Conference AS-SC. 2001 Consensus

        Guidelines for the management of women with cervical cytological abnormalities. JAMA.

        2002;287(16):2120-9.

39.   Katki HA, Schiffman M, Castle PE, Fetterman B, Poitras NE, Lorey T, et al. Five-year risks of CIN

        3+ and cervical cancer among women who test Pap-negative but are HPV-positive. J Low Genit

        Tract Dis. 2013;17(5 Suppl 1):S56-63.

40.   Moyer VA, Force USPST. Screening for lung cancer: U.S. Preventive Services Task Force

        recommendation statement. Ann Intern Med. 2014;160(5):330-8.

41.   Massad LS, Einstein MH, Huh WK, Katki HA, Kinney WK, Schiffman M, et al. 2012 updated

        consensus guidelines for the management of abnormal cervical cancer screening tests and

        cancer precursors. J Low Genit Tract Dis. 2013;17(5 Suppl 1):S1-S27.

42.   Wright TC, Jr., Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D, et al. 2006 consensus

        guidelines for the management of women with abnormal cervical cancer screening tests. Am J

        Obstet Gynecol. 2007;197(4):346-55.

43.   Perkins RB, Guido RS, Castle PE, Chelmow D, Einstein MH, Garcia F, et al. 2019 ASCCP Risk-Based

        Management Consensus Guidelines for Abnormal Cervical Cancer Screening Tests and Cancer

        Precursors. Journal of Lower Genital Tract Disease. 2020;24(2):102-31.

44.   Hedman E, Thompson P, Hansen K. Compliance With ASCCP Guidelines for Evaluation and

Management of Abnormal Pap Smears [28E]. Obstetrics & Gynecology. 2017;129(5):59S.

45.   Cheung LC, Egemen D, Chen X, Katki HA, Demarco M, Wiser AL, et al. 2019 ASCCP Risk-Based

Management Consensus Guidelines: Methods for Risk Estimation, Recommended Management,

and Validation. Journal of Lower Genital Tract Disease. 2020;24(2):90-101.

46.   Egemen D, Cheung LC, Chen X, Demarco M, Perkins RB, Kinney W, et al. Risk Estimates

Supporting the 2019 ASCCP Risk-Based Management Consensus Guidelines. Journal of Lower

Genital Tract Disease. 2020;24(2):132-43.

47.   Rendle KA, Schiffman M, Cheung LC, Kinney WK, Fetterman B, Poitras NE, et al. Adherence

patterns to extended cervical screening intervals in women undergoing human papillomavirus

(HPV) and cytology cotesting. Prev Med. 2018;109:44-50.

48.   Kinney W, Hunt WC, Dinkelspiel H, Robertson M, Cuzick J, Wheeler CM, et al. Cervical excisional

treatment of young women: a population-based study. Gynecol Oncol. 2014;132(3):628-35.

49.   Cuzick J, Myers O, Hunt WC, Saslow D, Castle PE, Kinney W, et al. Human papillomavirus testing

2007-2012: co-testing and triage utilization and impact on subsequent clinical management. Int

J Cancer. 2015;136(12):2854-63.

50.   Kasiske BL, Snyder JJ, Gilbertson D, Matas AJ. Diabetes mellitus after kidney transplantation in

the United States. Am J Transplant. 2003;3(2):178-85.

51.   Kasiske BL, Chakkera HA, Roel J. Explained and unexplained ischemic heart disease risk after

renal transplantation. J Am Soc Nephrol. 2000;11(9):1735-43.

52.   Davidson J, Wilkinson A, Dantal J, Dotta F, Haller H, Hernandez D, et al. New-onset diabetes after

transplantation: 2003 International consensus guidelines. Proceedings of an international expert

panel meeting. Barcelona, Spain, 19 February 2003. Transplantation. 2003;75(10 Suppl):SS3-24.

53. Baid S, Cosimi AB, Farrell ML, Schoenfeld DA, Feng S, Chung RT, et al. Posttransplant diabetes mellitus in liver transplant recipients: risk factors, temporal relationship with hepatitis C virus allograft hepatitis, and impact on mortality. Transplantation. 2001;72(6):1066-72.

54. Knobler H, Stagnaro-Green A, Wallenstein S, Schwartz M, Roman SH. Higher incidence of diabetes in liver transplant recipients with hepatitis C. J Clin Gastroenterol. 1998;26(1):30-3.

55. Ye X, Kuo HT, Sampaio MS, Jiang Y, Bunnapradist S. Risk factors for development of new-onset diabetes mellitus after transplant in adult lung transplant recipients. Clin Transplant. 2011;25(6):885-91.

56. Pham PT, Pham PM, Pham SV, Pham PA, Pham PC. New onset diabetes after transplantation (NODAT): an overview. Diabetes Metab Syndr Obes. 2011;4:175-86.

57. UNOS 2019;Pageshttps://unos.org/data/transplant-trends/ on 04/15/2019.

58. Sharif A, Hecking M, de Vries AP, Porrini E, Hornum M, Rasoul-Rockenschaub S, et al. Proceedings from an international consensus meeting on posttransplantation diabetes mellitus: recommendations and future directions. Am J Transplant. 2014;14(9):1992-2000.

59. Munshi VN, Saghafian S, Werner K, Cook CB, Chakkera HA. Comparison of Post-Transplantation Diabetes Mellitus Incidence and Risk Factors between Kidney and Liver Transplantation Patients. Submitted; 2019.

60. Munshi VN, Saghafian S, Cook CB, Eric Steidley D, Hardaway B, Chakkera HA. Incidence, Risk Factors, and Trends for Post-Heart Transplantation Diabetes Mellitus. American Journal of Cardiology.

61. Boloori A, Saghafian S, Chakkera HA, Cook CB. Characterization of Remitting and Relapsing Hyperglycemia in Post-Renal-Transplant Recipients. PLoS One. 2015;10(11):e0142363.

62. Chakkera HA, Pham PT, Pomeroy J, Weil EJ, Knowler WC. Response to Comment on: Chakkera et al. Can New-Onset Diabetes After Kidney Transplant Be Prevented? Diabetes Care 2013;36:1406-1412. Diabetes Care. 2013;36(10):e183.

63. Xu S, Schroeder E, Shetterly S, Goodrich G, O'Connor P, Steiner J, et al. Accuracy of hemoglobin A1c imputation using fasting plasma glucose in diabetes research using electronic health records data. Statistics, Optimization & Information Computing. 2014;2.

64. van der Heijden GJ, Donders AR, Stijnen T, Moons KG. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. J Clin Epidemiol. 2006;59(10):1102-9.

65. Janssen KJ, Donders AR, Harrell FE, Jr., Vergouwe Y, Chen Q, Grobbee DE, et al. Missing covariate data in medical research: to impute is better than to ignore. J Clin Epidemiol. 2010;63(7):721-7.

66. Hung AM, Roumie CL, Greevy RA, Liu X, Grijalva CG, Murff HJ, et al. Kidney function decline in metformin versus sulfonylurea initiators: assessment of time-dependent contribution of weight, blood pressure, and glycemic control. Pharmacoepidemiol Drug Saf. 2013;22(6):623-31.

67. Masica AL, Ewen E, Daoud YA, Cheng D, Franceschini N, Kudyakov RE, et al. Comparative effectiveness research using electronic health records: impacts of oral antidiabetic drugs on the development of chronic kidney disease. Pharmacoepidemiol Drug Saf. 2013;22(4):413-22.

68. Ramos M, McEwan P, Lamotte M, Foos V. The Relationship Of Predicted Benefits In Life Expectancy And Quality Adjusted Life Expectancy For Improved Glucose Control In Type 2 Diabetes Simulation Modeling. Value in Health. 2017;20(9):A747-A8.

69. Zhou H, Isaman DJM, Messinger S, Brown MB, Klein R, Brandle M, et al. A Computer Simulation Model of Diabetes Progression, Quality of Life, and Cost. Diabetes Care. 2005;28(12):2856.

70. Mount Hood 4 Modeling G. Computer modeling of diabetes and its complications: a report on the Fourth Mount Hood Challenge Meeting. Diabetes Care. 2007;30(6):1638-46.

71. Leal J, Gray AM, Clarke PM**.** Development of life-expectancy tables for people with type 2 diabetes. Eur Heart J. 2009;30(7):834-9.

72. Chakkera HA, Knowler WC, Devarapalli Y, Weil EJ, Heilman RL, Dueck A, et al. Relationship between inpatient hyperglycemia and insulin treatment after kidney transplantation and future new onset diabetes mellitus. Clin J Am Soc Nephrol. 2010;5(9):1669-75.

73. Chakkera HA, Weil EJ, Castro J, Heilman RL, Reddy KS, Mazur MJ, et al. Hyperglycemia during the immediate period after kidney transplantation. Clin J Am Soc Nephrol. 2009;4(4):853-9.

74. Rao PS, Schaubel DE, Jia X, Li S, Port FK, Saran R**.** Survival on dialysis post-kidney transplant failure: results from the Scientific Registry of Transplant Recipients. Am J Kidney Dis. 2007;49(2):294-300.

75. Gill JS, Abichandani R, Kausz AT, Pereira BJ**.** Mortality after kidney transplant failure: the impact of non-immunologic factors. Kidney Int. 2002;62(5):1875-83.

76. Cosio FG, Pesavento TE, Kim S, Osei K, Henry M, Ferguson RM**.** Patient survival after renal transplantation: IV. Impact of post-transplant diabetes. Kidney Int. 2002;62(4):1440-6.

77. Shivaswamy V, Boerner B, Larsen J**.** Post-Transplant Diabetes Mellitus: Causes, Treatment, and Impact on Outcomes. Endocr Rev. 2016;37(1):37-61.

78. Arias E, Xu J**.** United States Life Tables, 2015. Natl Vital Stat Rep. 2018;67(7):1-64.

79. Kaballo MA, Canney M, O'Kelly P, Williams Y, O'Seaghdha CM, Conlon PJ**.** A comparative analysis of survival of patients on dialysis and after kidney transplantation. Clin Kidney J. 2018;11(3):389-93.

80. Morales JM, Marcen R, del Castillo D, Andres A, Gonzalez-Molina M, Oppenheimer F, et al. Risk factors for graft loss and mortality after renal transplantation according to recipient age: a prospective multicentre study. Nephrol Dial Transplant. 2012;27 Suppl 4:iv39-46.

81.   Wong RJ, Cheung R, Perumpail RB, Holt EW, Ahmed A**.** Diabetes mellitus, and not obesity, is

associated with lower survival following liver transplantation. Dig Dis Sci. 2015;60(4):1036-44.

82.   Deo SV, Sarabu N, Kumar S, Altarabsheh S, Dunlay SM, Kilic A, et al. Abstract 12017: New-Onset

Diabetes After Heart Transplantation is associated With Worse Long-term Survival: A Propensity

Matched Analysis of a National Registry. Circulation. 2016;134(suppl_1):A12017-A.