Curriculum Vitae

ANNE-MICHELLE NOONE
annie.noone@gmail.com

EDUCATION

University of Maryland Baltimore, Baltimore, Maryland Spring 2021
Doctor of Philosophy, Epidemiology
Mentors: Eric A. Engels, M.D., MPH and Joanne F. Dorgan, Ph.D.

University of Michigan, Ann Arbor, Michigan Spring 2005
Master of Science, Biostatistics

Boston University, Boston, Massachusetts Spring 2003
Master of Arts & Bachelor of Arts, Mathematics
Specialty in Statistics
Minors: French & Philosophy

PROFESSIONAL EXPERIENCE

Mathematical Statistician July 2010 – present
Data Analytics Branch
Surveillance Research Program, Division of Cancer Control and Population Sciences,
National Cancer Institute, Rockville, Maryland

- Lead planning, evaluation, and production of statistical programs and reports.
  - Lead annual preparation of Surveillance, Epidemiology and End Results (SEER) data for publication. Activities include a quality assessment of variables and annual publication of rates and trends in cancer incidence, survival and mortality.
  - Evaluate trends in population-based cancer statistics and their relationship to determinants of cancer incidence, survival, or mortality.
  - Led evaluation and validation of treatment data collected by population-based cancer registries using multiple sources such as Medicare claims, Patterns of Care special studies and external data sources.
- Develop novel statistical methods applicable to population-based cancer registry data.
  - Applied machine learning techniques to cancer registry data linked to Medicare claims in a novel method to identify cases of bladder cancer missed by the cancer registries.
  - Developed more representative populations for the estimation of endometrial cancer incidence rates. These populations led to more accurate reporting of cancer incidence.

- Collaborate with and provide support to researchers within NCI and from the external research community to access and use cancer statistics and risk factors.
  - Participation on annual reports such as the Annual Report to the Nation on the Status of Cancer which is a collaboration with several partner agencies and the Cancer Trends Progress Report.
  - Served as the statistical and SEER data expert on various collaborative projects such as: Developing populations for subgroups of Asian and Pacific Islanders, analyzing and reporting cancer statistics using these special populations, and analyzing data from the Survey of Physician Attitudes Regarding the Care of Cancer Survivors (SPARCCS).
- Serve as the Level 3 Contracting Officer's Representative (COR) on the Division of Cancer Control and Population Sciences biomedical computing support contract.
  - Led the re-competition of this large and complex contract (total contract ceiling is $90M).
  - Leading development of internal processes to enhance efficiency and accountability from the contractor.
  - Primary COR responsible for monitoring and approving work, communicating with the contractor, and managing tasks given to the contractor.

Biostatistician April 2008 – July 2010
Lombardi Comprehensive Cancer Center (LCCC) &
Department of Biostatistics, Bioinformatics, and Biomathematics
Georgetown University Medical Center, Washington, District of Columbia

- Collaborate with LCCC investigators on basic science, clinical, and population science research projects, providing biostatistical expertise in assisting with the design, implementation, and analysis of a wide range of studies. Implement various modeling techniques such as Cox regression, generalized linear models (linear, logistic and Poisson regression), linear mixed effects models, generalized estimating equations to account for clustering by study site and multiple imputation to handle missing data
- Assist in the design and planning of institutional clinical trials. As the trial biostatistician, responsibilities include preparing monthly data monitoring reports for review by the Data Safety and Monitoring Committee and conducting statistical analysis.
- Provide biostatistical reviews of proposed clinical trials in the LCCC by participating on the Clinical Research Committee.
- Monitor and provide biostatistical reviews of on-going clinical trials through membership on the Data Safety and Monitoring Committee.
- Assist in the development and preparation of grant proposals.

TEACHING EXPERIENCE

Department of Epidemiology
University of Maryland Baltimore, Baltimore, Maryland

- Teaching Assistant (Fall 2015); Biostatistical Methods
- Teaching Assistant (Spring 2016); Principles of Epidemiology

Department of Biostatistics, Bioinformatics, and Biomathematics
Georgetown University Medical Center, Washington, District of Columbia

- Instructor (Fall 2008 & 2009), Course Director (Fall 2008); Introduction to Computational Software, taught unit on SAS programming.
- Instructor (Spring 2009); Pattern Recognition, taught unit on classification and regression trees (CART) and random forests.

COURSEWORK & TECHNICAL SKILLS

- Additional Coursework: Analysis of clinical trials, Bayesian methods, Longitudinal data analysis
- Epidemiologic Applications of Geographical Information Systems, The Johns Hopkins Bloomberg School of Public Health 2009 Graduate Summer Institute of Epidemiology and Biostatistics
- SEER-Medicare Data Training Workshop, National Cancer Institute, March 2010
- Experienced with SAS and R statistical software
- Experience with LaTeX, Microsoft Word, Excel and PowerPoint

AWARDS & HONORS

- Journal of Registry Management 2015 Best Paper Award
- NIH Cancer Research Training Grant, UM Department of Biostatistics (2003 – 2006)
- University Scholar, Boston University (1998 – 2002)

PUBLICATIONS IN REFEREED JOURNALS

1. Forjaz G, Bastos J, Castro C, Mayer A, **Noone AM**, et al. Regional differences in tobacco smoking and lung cancer in Portugal in 2018: a population-based analysis using nationwide incidence and mortality data. *BMJ Open* 10(10), 2020
2. Rivera DR, Lam CJK, Enewold L, Petkov VI, Tran Q, Brennan S, Dickie L, McNeel TS, **Noone AM**, et al. Development and utility of the observational research in oncology toolbox: Cancer medications enquiry database-Healthcare Common Procedure Coding System (HCPCS). *J Natl Cancer Inst Monogr* 55:39-45, 2020

3. **Noone AM**, Pfeiffer RM, Dorgan JF, et al. Cancer-Attributable Mortality among Solid Organ Transplant Recipients in the United States, 1987-2014. *Cancer* 125(15): 2647-2655, 2019
4. Temkin SM, Kohn EC, Penberthy L, Cronin KA, Rubisak L, Dickie LA, Minasian L**,** **Noone AM**. Hysterectomy-corrected rates of endometrial cancer among women younger than age 50 in the United States. *Cancer Causes Control* 29:427-433, 2018
5. Cronin KA, Lake AJ, Scott S, Sherman RL, **Noone AM**, et al. Annual Report to the Nation on the Status of Cancer, part I: National cancer statistics. *Cancer*, 2018
6. **Noone AM**, Negoita S, Schussler N, et al. Medical Record-Documented TNM Categories and Stage Group: Feasibility of Use for Cancer Surveillance. *J Registry Manag* 44:46-53, 2017
7. **Noone AM**, Cronin KA, Altekruse SF, et al: Cancer Incidence and Survival Trends by Subtype Using Data from the Surveillance Epidemiology and End Results Program, 1992-2013. *Cancer Epidemiol Biomarkers Prev* 26:632-641, 2017
8. Andridge R, **Noone AM**, Howlader N. Imputing estrogen receptor (ER) status in a population-based cancer registry: a sensitivity analysis. *Stat Med* 36:1014-1028, 2017
9. Temkin SM, Minasian L, **Noone AM**. The End of the Hysterectomy Epidemic and Endometrial Cancer Incidence: What Are the Unintended Consequences of Declining Hysterectomy Rates? Front Oncol. 2016 Apr 14;6:89.
10. Ryerson AB, Eheman CR, Altekruse SF, Ward JW, Jemal A, Sherman RL, Henley SJ, Holtzman D, Lake A, **Noone AM**, Anderson RN, Ma J, Ly KN, Cronin KA, Penberthy L, Kohler BA. Annual Report to the Nation on the Status of Cancer, 1975-2012, featuring the increasing incidence of liver cancer. *Cancer*. 2016 May 1;122(9):1312-37.
11. **Noone AM**, Schussler N, Negoita S, Adamo P, Cronin K, Cyr J, Gress D, Kosary C, Liu B, Sun L, Ward K, Penberthy L. Availability of TNM Staging Data Elements in the Medical Record and Training Needs Assessment: Results from the 2014 SEER Training Needs Assessment for TNM Study. *Journal of Registry Management,* 2015 Summer;42(2):40-7.
12. Kohler BA, Sherman RL, Howlader N, Jemal A, Ryerson AB, Henry KA, Boscoe FP, Cronin KA, Lake A, **Noone AM**, Henley SJ, Eheman CR, Anderson RN, Penberthy L. Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. *Journal of the National Cancer Institute,* 2015; 107(6).
13. Warren JL, Butler EN, Stevens J, Lathan CS, **Noone AM**, Ward KC, Harlan LC. Receipt of chemotherapy among medicare patients with cancer by type of supplemental insurance. *Journal of Clinical Oncology,* 2015; 33(4): 312-8.
14. Mariotto AB, **Noone AM**, Howlader N, Cho H, Keel GE, Garshell J, Woloshin S, Schwartz LM. Cancer survival: an overview of measures, uses, and interpretation. *Journal National Cancer Institute Monograph*, 2014; 49: 145-86.

15. Jamison PM, Altekruse SF, Chang JT, Zahn J, Lee R, **Noone AM**, Barroilhet L. Site-specific factors for cancer of the corpus uteri from SEER registries: collaborative stage data collection system, version 1 and version 2. *Cancer,* 2014; Suppl 23: 3836-45.

16. Nguyen AB, Chawla N, **Noone AM**, Srinivasan S. Disaggregated data and beyond: future queries in cancer control research. *Cancer Epidemiology Biomarkers and Prevention,* 2014; 23(11): 2266-72.

17. Cheng I, Le GM, **Noone AM**, Gali K, Patel M, Haile RW, Wakelee HA, Gomez SL. Lung cancer incidence trends by histology type among Asian American, Native Hawaiian, and Pacific Islander populations in the United States, 1990-2010. *Cancer Epidemiology Biomarkers and Prevention,* 2014; 23(11): 2250-65.

18. **Noone AM**, Lund JL, Mariotto A, Cronin K, McNeel T, Deapen D, Warren JL. Comparison of SEER treatment tata with Medicare claims. *Medical Care*, 2014;

19. Edwards BK, **Noone AM**, Mariotto AB, Simard EP, Boscoe FP, Henley SJ, Jemal A, Cho H, Anderson RN, Kohler BA, Eheman CR, Ward EM. Annual Report to the Nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. *Cancer*, 2014; 120(9): 1290-314.

20. Liu L, **Noone AM**, Gomez SL, Scoppa S, Gibson JT, Lichtensztajn D, Fish K, Wilkens LR, Goodman MT, Morris C, Kwong S, Deapen D, Miller BA. Cancer incidence trends among native Hawaiians and Pacidic Islanders in the United States, 1990-2008. *Journal of the National Cancer Institite,* 2013; 105(15): 1086-95.

21. Gomez SL, **Noone AM**, Lichtensztajn DY, Scoppa S, Gibson JT, Liu L, Morris C, Kwong S, Fish K, Wilkens LR, Goodman MT, Deapen D, Miller BA. Cancer incidence trends among Asian American populations in the United States, 1990-2008. *Journal of the National Cancer Institute,* 2013; 105(15): 1096-110.

22. Cheung WY, Aziz N, **Noone AM**, Rowland JH, Potosky AL, Ayanian JZ, Virgo KS, Ganz PA, Stefanek M, Earle CC. Physician preferences and attitudes regarding different models of cancer survivorship care: a comparison of primary care providers and oncologists. *Journal of Cancer Survivivorship,* 2013;7:343-54.

23. Gomez SL, **Noone AM**, Lichtensztajn DY, Scoppa S, Gibson JT, Liu L, Morris C, Kwong S, Fish K, Wilkens LR, Goodman MT, Deapen D, Miller BA. Cancer incidence trends among Asian American populations in the United States, 1990-2008. *Journal of the National Cancer Institute,* 2013;105:1096-110.

24. Liu L, **Noone AM**, Gomez SL, Scoppa S, Gibson JT, Lichtensztajn D, Fish K, Wilkens LR, Goodman MT, Morris C, Kwong S, Deapen D, Miller BA. Cancer incidence trends among native Hawaiians and other Pacific Islanders in the United States, 1990-2008. *Journal of the National Cancer Institute,* 2013;105:1086-95.

25. Han PK, Klabunde CN, **Noone AM**, Earle CC, Ayanian JZ, Ganz PA, Virgo KS, Potosky AL. Physicians' beliefs about breast cancer surveillance testing are consistent with test overuse. *Medical Care,* 2013;51:315-23.

26. Jamison PM, **Noone AM**, Ries LA, Lee NC, Edwards BK. Trends in endometrial cancer incidence by race and histology with a correction for the prevalence of hysterectomy, SEER 1992 to 2008. *Cancer Epidemiology Biomarkers and Prevention,* 2013;22:233-41.

27. Jemal A, Simard EP, Dorell C, **Noone AM**, Markowitz LE, Kohler B, Eheman C, Saraiya M, Bandi P, Saslow D, Cronin KA, Watson M, Schiffman M, Henley SJ, Schymura MJ, Anderson RN, Yankey D, Edwards BK. Annual Report to the Nation on the Status of Cancer, 1975-2009, Featuring the burden and trends in human papillomavirus (HPV)-associated cancers and HPV vaccination Coverage Levels. *Journal of the National Cancer Institute,* 2013;105:175-201.

28. Selsky C, Luta G, **Noone AM**, Huerta EE, Mandelblatt JS. Internet access and online cancer information seeking among Latino immigrants from safety net clinics, *Journal of Health Communication,* 2013;18:58-70.

29. Banerjee M, Ding Y, **Noone AM**. Identifying representative trees from ensembles. *Statistics in Medicine,* 2012;31(15):1601-16.

30. Eheman C, Henley SJ, Ballard-Barbash R, Jacobs EJ, Schymura MJ, **Noone AM**, Pan L, Anderson RN, Fulton JE, Kohler BA, Jemal A, Ward E, Plescia M, Reis LA, Edwards BK. Annual Report to the Nation on the status of cancer, 1975-2008, featuring cancers associated with excess weight and lack of sufficient physical activity. *Cancer,* 2012;118:2338-66.

31. Howlader N, **Noone AM**, Yu M, Cronin KA. Use of imputed population-based cancer registry data as a method of accounting for missing information: application to estrogen receptor status for breast cancer. *American Journal of Epidemiology,* 2012;176:347-56.

32. Mueller NT, **Noone AM**, Luta G, Wallington SF, Huerta EE, Mandelblatt JS. Information channels associated with awareness of human papillomavirus infections and vaccination among Latino immigrants from safety net clinics. *Journal of Immigrant and Minority Health,* 2012;14:183-8.

33. Wallington SF, Luta G, **Noone AM**, Caicedo L, Lopez-Class M, Sheppard V, Spencer C, Mandelblatt J. Assessing the awareness of and willingness to participate in cancer clinical trials among immigrant latinos. *Journal of Community Health,* 2012;37:335-43.

34. Cavalli LR, **Noone AM**, Makambi KH, Rone JD, Kasid UN, Haddad BR. Loss of the BLID Tumor Suppressor Gene in Early-Onset Breast Cancer. *Cytogenetic and Genome Research*, 2011; 135(1):19-24.

35. MacLaughlin BW, Wang D, Liu N, **Noone AM**, Harazduk N, Lumpkin M, Haramati A, Saunders P, Dutton M, Amri H. Stress Biomarkers in Medical Students Participating in a Mind Body Medicine Skills Program. *Evidence Based Complementary and Alternative Medicine,* 2010. doi:10.1093/ecam/neq039.

36. Ahmadi N, Goldman R, Seillier-Moiseiwitsch F, **Noone AM**, Kosti O, Davidson BJ. Decreased risk of squamous cell carcinoma of the head and neck in users of

nonsteroidal anti-inflammatory drugs. *Internaltional Journal of Otolaryngology,* 2010; 424161.

37. Broustas CG, Ross JS, Yang Q, Sheehan CE, Riggins RB, Cavalli LR, Haddad BR, Seillier-Moiseiwitsch F, **Noone AM**, Kallakury BVS, Haffty BG, Clarke R, Kasid UN. The proapoptotic molecule BLID interacts with Bcl-XL and its downregulation in breast cancer correlates with poor disease-free and overall survival. *Clinical Cancer Research,* 2010;16:2939-48.

38. Dyba M, Wang A, **Noone AM**, Goerlitz D, Shields P, Zheng Y-L, Chung FL**.** Metabolism of isothiocyanates in individuals with positive and null GSTT1 and M1 genotypes after drinking watercress juice. *Clinical Nutrition,* 2010;29:813-8.

39. Mandelblatt JS, Sheppard VB, Hurria A, Kimmick G, Isaacs C, Taylor KL, Kornblith AB, **Noone AM**, Luta G, Tallarico M, Barry WT, Hunegs L, Zon R, Naughton M, Winer E, Hudis C, Edge SB, Cohen HJ, Muss H. Breast cancer adjuvant chemotherapy decisions in older women: the role of patient preference and interactions with physicians. *Journal of Clinical Oncology,* 2010;28:3146-53.

40. Sen A, Banerjee M, Li Y, **Noone AM**. A Bayesian approach to competing risks analysis with masked cause of death. *Statistics in Medicine,* 2010;29:1681-95.

41. Unger KR, Lominska CE, Deeken JF, Davidson BJ, Newkirk KA, Gagnon GJ, **Noone AM,** Slack RS, Harter KW. Fractionated stereotactic radiosurgery for reirradiation of head-and-neck cancer. *International Journal of Radiatition, Oncology, Biology and Physics,* 2010;77:1411-9.

42. Wycherly BJ, Berkowitz F, **Noone AM**, Kim HJ. Computed tomography and otosclerosis: a practical method to correlate the sites affected to hearing loss. *Annals of Otlogy, Rhinology and Laryngology,* 2010;119:789-94.

43. Aiello AE, Diez-Roux A, **Noone AM**, Ranjit N, Cushman M, Tsai MY, Szklo M. Socioeconomic and psychosocial gradients in cardiovascular pathogen burden and immune response: the multi-ethnic study of atherosclerosis. *Brain Behavior and Immunity,* 2009;23:663-71.

44. Liu MC, Shields PG, Warren RD, Cohen P, Wilkinson M, Ottaviano YL, Rao SB, Eng-Wong J, Seillier-Moiseiwitch F, **Noone AM**, Isaacs C. Circulating tumor cells: a useful predictor of treatment efficacy in metastatic breast cancer. *Journal of Clinical Oncology,* 2009;27:5153-9.

45. Wu RR, Lindenberg PA, Slack R, **Noone AM**, Marshall JL, He AR. Evaluation of hypertension as a marker of bevacizumab efficacy. *Journal Gastrointestinal Cancer,* 2009;40:101-8.

46. Yan B, **Noone AM**, Yee C, Banerjee M, Schwartz K, Simon MS. Racial differences in colorectal cancer survival in the Detroit Metropolitan Area. *Cancer,* 2009;115:3791-800.

47. Deng HY, Zwetchkenbaum S, **Noone AM**. Bond strength of silicone to polyurethane following immersion of silicone in cleaning solutions. *Journal of Prosthetic Dentistry,* 2004;91:582-5.

48. Kroon HM, Li Q, Teitz-Tennenbaum S, Whitfield JR, **Noone AM**, Chang AE. 4-1BB costimulation of effector T cells for adoptive immunotherapy of cancer: involvement of Bcl gene family members. *Journal of Immunotherapy,* 2007;30:406-16.
49. Movva S, **Noone AM,** Banerjee M, Patel D, Schwartz K, Yee CL, Simon MS. Racial differences in cervical cancer survival in the Detroit metropolitan area. *Cancer,* 2008;112:1264-71.
50. Simon MS, Banerjee M, Crossley-May H, Vigneau FD, **Noone AM**, Schwartz K. Racial differences in breast cancer survival in the Detroit Metropolitan area. *Breast Cancer Research and Treatment,* 2006;97:149-55.

PRESENTATIONS

- Cancer-attributable mortality among solid organ transplant recipients in the United States, 1987-2014, Society for Epidemiologic Research, 2019, Minneapolis, MN.
- Assigning AJCC TNM: results from a Field Study, National Cancer Registrar's Association Plenary Presentation, 2016, Las Vegas, NV.
- SEER 2014 Training Assessment for AJCC TNM Staging, National Cancer Registrar's Association Plenary Presentation, 2015, San Antonio, TX.
- Imputed population-based cancer registry data: Application to Estrogen Status (ER) for Breast Cancer. Lecturer in short course, NAACCR Annual Conference 2013, Austin, TX.
- Comparison of SEER Treatment Data using the Patterns of Care Studies. Presentation, NAACCR Annual Conference 2013, Austin, TX.
- Cancer incidence trends among Asian Americans, native Hawaiians and other Pacific Islanders in the United States, 1990-2008. Presentation, NAACCR Annual Conference 2013, Austin, TX.
- The Effect of Changing Hysterectomy Prevalence in Trends in Endometrial Cancer, SEER 1992-2008. Presentation, NAACCR Annual Conference 2012, Portland, OR.
- Validation of SEER Treatment Data Using Medicare Claims. Presentation, NAACCR Annual Conference 2012, Portland, OR.
- Multiple Primaries (MPs) in Survival Estimates: Should SEER include or exclude MPs? Poster, NAACCR Annual Conference 2011, Louisville, KY.
- Identifying Representative Trees in Random Forest for Survival Data. Presentation, Joint Statistical Meetings (JSM), Denver, Colorado, August 2008. Work with Mousumi Banerjee and Ying Ding from University of Michigan.

Abstract

Cancer Mortality among US Solid Organ Transplant Recipients:
Novel Methodologies to Estimate Cancer Burden using Linked Population-Based Registries

Anne-Michelle Noone, Doctor of Philosophy, 2021

Dissertation Directed By:

Joanne Dorgan PhD
Professor
Department of Epidemiology and Public Health

and

Eric A. Engels MD, MPH
Senior Investigator
Division of Cancer Epidemiology and Genetics, National Cancer Institute

Background: The solid organ transplant population has an elevated risk of cancer compared with the general population. Excess risk is largely due to immunosuppression. As this population grows, understanding long-term health risks such as cancer is critical. Population-based estimates of cancer mortality are needed since they measure the downstream outcome following a cancer diagnosis. Furthermore, quantifying deaths attributable to cancer can inform priorities to reduce the cancer burden.

Methods: Linked transplant and cancer registry data were used to identify incident cancers and deaths among solid organ transplant recipients in the United States (1987-2014). Population-attributable fractions (PAFs) of deaths due to cancer and corresponding cancer-attributable mortality rates were estimated. Cancer-attributable mortality rates computed using the PAF were compared to cancer-specific mortality rates computing using cause of death (COD). The life-years lost (LYL) to cancer were estimated using two methods: an approach using matching to construct a cancer-free cohort and an approach using Cox proportional hazards regression models.

Results: Among 221,962 transplant recipients, 15,012 developed cancer. Thirteen percent of deaths (PAF=13.2%) were attributable to cancer, corresponding to a cancer-attributable mortality rate of 516 per 100,000 person-years. Lung cancer was the largest contributor to mortality (PAF=3.1%), followed by non-Hodgkin lymphoma (NHL, PAF=1.9%), colorectal cancer (PAF=0.7%), and kidney cancer (PAF=0.5%). Overall, the cancer-specific mortality rate lower, 368 per 100,000 person-years.

Within 10 years post-transplant, the mean LYL was 0.16 years per transplant recipient and 2.7 years per cancer. Cancer accounted for 1.9% of the total LYL expected in this population. Lung cancer was the largest contributor, accounting for 24% of all LYL, and NHL had the next highest contribution (15%).

Conclusions: Cancer is a substantial cause of mortality among solid organ transplant recipients resulting in excess deaths and a shortened lifespan. Lung cancer and NHL are major contributors to the cancer burden including LYL to cancer, highlighting opportunities to reduce cancer mortality through prevention and screening.

Cancer Mortality among US Solid Organ Transplant Recipients: Novel Methodologies to Estimate Cancer Burden using Linked Population-Based Registries

by

Anne-Michelle Noone

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, Baltimore in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Dedication

To Helen – Anything is possible

Acknowledgements

I would like to thank Drs. Eric Engels and Joanne Dorgan for their mentorship. I am extremely grateful for their support, patience, numerous discussions and edits throughout this dissertation. I am very grateful to Drs. Larry Magder and Ruth Pfeiffer for their statistical mentoring and insights to this analysis, and Dr. Jonathan Bromberg for his insights into the clinical relevance of this work. I am grateful to have a dissertation that perfectly combines my interest in advancing population-based epidemiology with applied biostatical methods.

I also acknowledge the transplant and cancer registries who participated in the Transplant Cancer Match study that created the basis for this dissertation.

Thank you to my colleagues and friends at the Surveillance Research Program of the National Cancer Institute who offered advice, support, friendship, and coffee breaks. A special thank you to Angela Mariotto, Lynne Penberthy, Kathy Cronin, and Nadia Howlader.

And finally, thank you to my family for their love and support. Markus, thank you for always listening to and supporting my wild ideas. Helen, I love you, lovie! Mom, those hot fudge fractions really paid off. Dad, thanks for seeing me through.

During the journey of this dissertation I have had the fortune of many friends touching my life and I thank you for the love and lessons you have given.

Table of Contents

Chapter Page

Dedication …………………………………………………………………………………iii

1. INTRODUCTION AND SPECIFIC AIMS... 1

2. BACKGROUND 4

   A. Solid organ transplant recipients... 4

   B. Risk factors for cancer among solid organ transplant recipients 6

   C. Epidemiology of cancer among solid organ transplant recipients 8

   D. Metrics to quantify the cancer mortality burden.. 9

   E. Methods to quantify the cancer mortality burden 11

   F. Significance.. 17

   G. Innovation 19

3. STUDY DESIGN AND METHODS.. 21

   A. Data source for all aims: Transplant Cancer Match study: A linkage between US transplant registry and cancer registries 21

   B. Aim 1: Estimate proportion of deaths due to cancer and the cancer-attributable mortality rates among transplant recipients using two statistical approaches 22

      1) Aim 1.1: Estimate proportion of deaths associated with cancer and the cancer-attributable mortality rate using population attributable fraction methodology 23

      2) Aim 1.2: Estimate the proportion of deaths due to cancer and the cancer-attributable mortality rate using COD from the cancer registries.. 31

      3) Aim 1.3: Compare estimates from aim 1.1 and 1.2 33

   C. Aim 2: Estimate mean LYL due to cancer among transplant recipients .. 34

      1) Aim 2.1: Estimate mean LYL due to cancer using matching methods... 35

2) Aim 2.2: Estimate the mean LYL using modeling. Cox models were used to estimate the counterfactual, which is survival had no cancer occurred among transplant recipients. 40

3) Aim 2.3: Compare LYL estimates from both methods................ 46

4. CANCER-ATTRIBUTABLE MORTALITY AMONG SOLID ORGAN TRANSPLANT RECIPIENTS IN THE UNITED STATES, 1987 THROUGH 2014.. 47

A. Abstract 47

B. Introduction 48

C. Methods... 49

D. Results 52

E. Discussion 70

5. CANCER-ATTRIBUTABLE MORTALITY USING CAUSE OF DEATH AND COMPARISON OF RESULTS (AIM 1.2 AND 1.3).. 77

A. Potential impact of unmeasured confounding on the hazard ratio used to calculate the population attributable fraction... 77

1) Estimate what the hazard ratio would need to be so that the cancer-attributable mortality rate from the PAF method equals that from the COD method 77

B. Coding of cause of death.. 84

1) Comparison of cause of death from SRTR and cancer registries . 84

2) Fraction of deaths from other causes attributable to cancer 85

C. Discussion 87

6. LIFE-YEARS LOST TO CANCER AMONG SOLID ORGAN TRANSPLANT RECIPIENTS 88

A. Abstract 88

B. Introduction 89

C. Methods... 90

D. Results 102

E. Discussion ... 109

7. LIFE-YEARS LOST DUE TO CANCER ESTIMATED USING MATCHING AND COMPARISON OF RESULTS (AIM 2.1 AND 2.3) ... 113

A. Introduction ... 113

B. Methods ... 113

C. Results ... 118

D. Comparison and discussion of methods ... 133

8. DISCUSSION ... 136

A. Risk factors for cancer ... 136

B. Cancer incidence and mortality in the transplant population ... 137

C. Statistical methods ... 140

D. Population-attributable fraction and cancer-attributable mortality rate ... 145

E. Cancer-attributable mortality using cause of death ... 147

F. Life-years lost to cancer ... 149

G. Life-years lost to cancer using a matching method ... 150

H. Strengths and limitations ... 152

I. Prevention and screening ... 153

J. Conclusion ... 154

APPENDIX 1: STATISTICAL CONSIDERATIONS IN ESTIMATING THE POPULATION-ATTRIBUTABLE FRACTION ... 155

A. Modeling of age in the Cox regression model used to estimate the population-attributable fraction ... 155

B. Estimate of the population-attributable fraction in the presence of an interaction between cancer and follow-up time in the Cox model ... 160

C. Validity of proportional hazards assumption for the Cox model used to estimate the population-attributable fraction ... 167

D. Sensitivity of the PAF to exclusions of cancers within one year after transplant ... 185

APPENDIX 2: STATISTICAL CONSIDERATIONS IN ESTIMATING THE LIFE-YEARS LOST TO CANCER 188

A. Selection of censoring at 10 years post-transplant 188

B. Functional form of continuous covariates 189

C. Selection of caliper width for matching 190

REFERENCES 193

LIST OF TABLES

Table 1: Estimated population attributable fraction and its 95% confidence interval for two cohorts 29

Table 2: Characteristics of 221,962 solid organ transplant recipients in the United States (1987-2014) 53

Table 3: Cancer-attributable mortality among solid organ transplant recipients 54

Table 4: Cancer-attributable mortality rates, according to individual and transplant characteristics and cancer site 55

Table 5: PAF and cancer-attributable mortality for five mortality outcomes 69

Table 6: Parameter estimates used in bias factor analysis 76

Table 7: Comparison of mortality rates estimated using COD and PAF from registries who provided COD information 80

Table 8: Hazard ratio of death adjusted for potential confounders by cancer type 82

Table 9: Distribution of COD from SRTR among those with incident cancer and those without 85

Table 10: Comparison of COD from SRTR to COD from cancer registries among those with cancer 85

Table 11: Population attributable fraction and cancer-attributable mortality for other causes of death among 165,839 recipients with COD from cancer registries ... 86

Table 12: Characteristics of transplant recipients in the United States, by transplanted organ 92

Table 13: Additional characteristics of transplant recipients by organ transplanted 94

Table 14: Hazard ratios from the Cox proportional hazards models used to estimate the overall survival and cancer-free survival experience for kidney and liver recipients 97

Table 15: Hazard ratios from the Cox proportional hazards models used to estimate the overall survival and cancer-free survival experience for heart and lung recipients 99

Table 16: Hazard ratios from the Cox proportional hazards models used to estimate the overall survival and cancer-free survival experience for recipients of other or multiple organs 101

Table 17: Life-years lost to cancer among US solid organ transplant recipients 103

Table 18: Life-years to cancer among US solid organ transplant recipients, by cancer site 106

Table 19: Comparison of estimated life-years lost to cancer among solid organ transplant recipients using two approaches, one based on survival curves estimated from the Cox proportional hazards model and one based on Kaplan Meier (KM) survival curves 108

Table 20: Example LYL calculation using matching with time stratified by time since diagnosis/case selection for kidney recipients 117

Table 21: Pooled estimates from the Cox model using imputed data for kidney and liver transplant recipients 118

Table 22: Pooled estimates from the Cox model using imputed data for heart and lung transplant recipients 120

Table 23: Pooled estimates from the Cox model using imputed data for transplant recipients of other organ types 121

Table 24: Number of cancer cases and available matches within caliper width of 0.2 .. 124

Table 25: Comparison of characteristics of cases and controls for kidney and liver recipients based on top 5 matches and caliper width of 0.2 125

Table 26: Comparison of characteristics of cases and controls for heart and lung based on top 5 matches and caliper width of 0.2 127

Table 27: Comparison of characteristics of cases and controls for recipients of other organs based on top 5 matches and caliper width of 0.2 129

Table 28: Life-years lost to cancer by organ using the matching method with stratified Kaplan-Meier curves 130

Table 29: Comparison of life-years lost to cancer from each method after matching 131

Table 30: Comparison of life-years lost from matching controls on prognostic vs. propensity scores 133

Table 31: Comparison of the total LYL and proportion of LYL from the selected matching method and modeling methods 135

Table 32: Summary of AIC for Cox models with different functional forms for age 159

Table 33: Estimated effect of cancer with different functional forms for age 159

Table 34: Comparison of PAF estimates from 3 methods 165

Table 35: Relative risk of death among recipients by sex 175

Table 36: Relative risk of death among recipients by transplanted organ 176

Table 37: Relative risk of death among recipients by age at transplant 177

Table 38: Relative risk of death among recipients by transplant year 178

Table 39: Relative risk of death by transplant year with all recipients censored at 3 years post-transplant 181

Table 40: Hazard ratios for associations with death among recipients with an interaction between cancer and time since cancer diagnosis 182

Table 41: Hazard ratios for associations with death among recipients with an interaction between cancer and collapsed categories for time since cancer diagnosis 183

Table 42: Relative risk of death among recipients with interaction between cancer effect and follow-up time 184

Table 43: Population-attributable fraction and mortality rates after excluding cancers 90 days and 1 year post-transplant 186

Table 44: Deaths at each time interval with censoring at cancer incidence 189

Table 45: Number of controls per case for a range of caliper widths 191

LIST OF FIGURES

Figure 1: Population attributable fraction calculation 25

Figure 2: Life-years lost due to cancer among transplant recipients 41

Figure 3: Cancer-attributable mortality rates, according to age at transplant................... 62

Figure 4: Cancer-attributable mortality rates, according to sex.. 63

Figure 5: Cancer-attributable mortality rates, according to race/ethnicity 65

Figure 6: Cancer-attributable mortality rates, according to transplanted organ 66

Figure 7: Cancer-attributable mortality rates, according to time since transplant............ 68

Figure 8: Kaplan-Meier curves for the cohort of transplant recipients under actual and hypothetical cancer-free scenarios 105

Figure 9: Proportion of expected life-years that were lost to cancer by age and transplanted organ 109

Figure 10: Kaplan-Meier curves for cancer cases stratified by time since transplant 116

Figure 11: Kaplan-Meier curves for matched controls stratified by time since selection 116

Figure 12: Estimates of prognostic score by cancer status in whole population before matching 122

Figure 13: Graphical representation of the population attributable fraction among those with cancer 141

Figure 14: Schematic showing estimation of life-years lost... 143

Figure 15: Life-years lost due to cancer among transplant recipients 144

Figure 16: Estimated log-hazard ratios for age as a categorical variable from the Cox model used to estimate the population-attributable fraction 155

Figure 17: Estimated spline transformation of age from the Cox model used to estimate the population-attributable fraction 156

Figure 18: Estimated log-hazard ratios from Cox model stratified by age and then age as categorical within strata... 157

Figure 19: Estimated log hazard ratio association from model with age as a piecewise linear spline controlling for cancer 158

Figure 20: Survival plots by race 168

Figure 21: Survival plots by sex 168

Figure 22: Survival plots by year of transplant 169

Figure 23: Survival plots by transplanted organ 169

Figure 24: Survival plots by transplanted organ with collapsed groupings 170

Figure 25: Schoenfeld residuals with loess for sex 171

Figure 26: Schoenfeld residuals with loess by race (Non-Hispanic white is reference group) 172

Figure 27: Schoenfeld residuals with loess by organ (kidney is reference group) 173

Figure 28: Schoenfeld residuals with loess by year of transplant (1987-1994 is reference group) 174

Figure 29: Distribution of follow-up time among recipients who developed incident cancer and those who did not 188

Figure 30: Estimates log hazard ratios for years from initial dialysis to transplant among kidney recipients 190

LIST OF ABBREVIATIONS

AIDS	acquired immunodeficiency syndrome
API	Asian/Pacific islander
BMI	body mass index
CI	confidence interval
COD	cause of death
COPD	chronic obstructive pulmonary disease
CRC	colorectal cancer
CT	computerized tomography
EBV	Epstein-Barr virus
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HR	hazard ratio
IPF	idiopathic pulmonary fibrosis
KM	Kaplan-Meier
LYL	life-years lost
mTOR	mammalian target of rapamycin
NH	non-Hispanic
NHL	non-Hodgkin lymphoma
PAF	population attributable fraction
PBC	primary biliary cirrhosis
PSC	primary sclerosing cirrhosis

PTLD	post-transplant lymphoproliferative disorder
PY	person-years
RR	relative risk
SD	standard deviation
SE	standard error
SEER	Surveillance, Epidemiology, and End Results
SMR	standardized mortality ratio
SRTR	Scientific Registry of Transplant Recipients
TCM	Transplant Cancer Match
UK	United Kingdom
US	United States
UV	ultraviolet

1. Introduction and specific aims

Solid-organ transplant recipients have a two-fold increased risk for developing cancer compared with the general population.[1] The excess risk varies by cancer type and is largely due to immunosuppression and viral infections. In addition to cancer incidence, cancer mortality is an important measure of burden since it reflects the final downstream outcome in a population accounting for cancer incidence, treatment, and survival following a cancer diagnosis. Population-based mortality estimates and trends quantify the magnitude of the cancer burden in the transplant population and can be used to monitor changes over time. Thus, they provide valuable information for identifying subgroups that may benefit from screening or lifestyle interventions aimed at cancer prevention. Several studies outside of the US have shown an increased risk of death from cancer among transplant recipients.[2-5] Only one study of cancer mortality in this population has been conducted in the US and it had substantial opportunity for bias and only included kidney recipients. Also, no study has estimated cancer mortality over time, which is likely to change along with organ allocation procedures, cancer detection and treatment in this population.

Because accurate estimates of cancer mortality rates for the US transplant population are not currently available, this dissertation addresses a gap in knowledge. The overall objective is to quantify the burden of cancer mortality among US solid organ transplant recipients using novel approaches to estimate mortality statistics that reflect different aspects of the mortality burden. Three dimensions of the cancer mortality burden are estimated: the proportion of deaths due to cancer, the cancer-attributable mortality rate, and the life-years lost due to cancer.

The proportion of deaths attributed to cancer and cancer-attributable mortality rates were estimated using two approaches. The first used the population attributable fraction (PAF). An advantage of this method is that it does not rely on cause of death information collected on transplant recipients that may be unreliable. The second approach estimated these statistics using cause of death information from cancer registries. The results from this method were compared to results using the PAF methodology.

I then estimated the expected life-years lost (LYL) due to cancer among transplant recipients. In contrast to the mortality rate, LYL inherently incorporates age at death since younger people have longer potential future lifetimes. Thus, LYL can help assess the social and economic impact associated with cancer rather than simply the number of deaths occurring in the population. I used methods that are not widely used in cancer epidemiology and have never been applied to the transplant population. The aims below describe two methodologic approaches to estimate each mortality measure and a comparison of those estimates.

Analyses were conducted using data from the National Cancer Institute's Transplant Cancer Match Study, which is a linkage of the US Scientific Registry of Transplant Recipients (1987-2013) and 17 population-based state and regional cancer registries. The transplant registry contains data on more than 290,000 solid organ recipients who live in the cancer registry catchment areas and comprises approximately 50% of the US total.

Specific aims

Aim 1: Estimate proportion of deaths due to cancer and the cancer-attributable mortality rates among transplant recipients using:

Aim 1.1: Population attributable fraction methodology

Aim 1.2: Cause of death (COD) from cancer registries

Aim 1.3: Comparison of estimates from aim 1.1 and 1.2

Aim 2: Estimate mean life-years lost due to cancer among transplant recipients by:

Aim 2.1: Approximating the counterfactual life years lived without cancer by matching to recipients without cancer

Aim 2.2: Modeling the survival distributions

Aim 2.3: Comparison of estimates from aim 2.1 and 2.2

2. Background

A. Solid organ transplant recipients

Solid organ transplantation is a life-saving procedure performed for those with end-stage organ disease. The number of solid organ transplants has been increasing over time due to increasing need and better surgical procedures. In 2016, 33,611 solid organ transplants were performed in the United States, almost triple the number performed in 1988.[6] The most common transplanted organ was kidney with 19,061 transplants performed. This was followed by 7,841 liver transplants, 3,191 heart transplants, and 2,327 lung transplants. Organ allocation procedures are organ specific. In general, kidneys are allocated based on waiting time and expected survival benefit. Livers and hearts are allocated based on medical need and lungs are allocated based on survival benefit, medical urgency and waiting time.[7]

In 2016, the majority of US transplant recipients were white and a majority were male. Over time, the age at transplant has been increasing. In 1988, 63% of recipients were age 35 or older and by 2016, this age group comprised 83% of recipients.[6] In particular, the proportion of transplant recipients age 65 or older has been increasing. In 1988, 2% of recipients were age 65 or older and by 2016, this age group comprised over 18%.[6] The distribution of organs transplanted has also changed over time. In 1988, 70% of transplants were kidney, followed by liver and heart (about 14% each). In 2016, liver transplants increased to comprise 23% of all transplanted organs, heart transplants decreased to 9%. The largest change was that 7% of transplanted organs were lung in 2016 compared to only 0.3% in 1988.

In order to prevent rejection of the transplanted organ, recipients must take immunosuppressant medications. These medications are used to prevent rejection of the organ, as induction at the time of transplant, for maintenance treatment over an extended time, and acutely to treat the development of rejection. There are four classes of maintenance drugs: calcinuerin inhibitors, antiproliferative agents, mTOR inhibitors, and steroids. The specific regimen is tailored to the patient and the combinations used have changed over time. In the early 1990s, cyclosporine was the most commonly used calcinuerin inhibitor, used by over 90% of recipients. However, its usage declined in favor of tacrolimus. In recent years, tacrolimus was used by 80% or more recipients.[8-12] Similarly, azathioprine was the most commonly used antimetabolite in the early 1990s but usage started to decline in the mid-1990s in favor of mycophenolate.[8-12]

In addition to the increasing number of transplants performed, survival of the graft and patient has improved over time. In 2010, over 90% of kidney recipients were alive five years after transplant.[10] The five-year survival rate was 74% for liver recipients and 77% for heart recipients.[9,11] Lung recipients had the lowest five-year survival at 56%.[8] The improved survival rates are a result of better medical management including use of more efficacious immunosuppressant drugs and changes in monitoring and treatment for rejection.

Indeed, the increased number of transplants, improved outcomes and survival have contributed to a growing population of transplant recipients. In 2012, the population of transplant recipients in the US was approximately 282,200 people which was more than double the size in 1988.[13] Since the population of transplant recipients is increasing, understanding the long-term health risks in solid organ transplant recipients is important.

B. Risk factors for cancer among solid organ transplant recipients

As survival of transplant recipients has improved over time, cancer has become one of the leading causes of morbidity and mortality in this population. Transplant recipients may have health behaviors known to increase cancer risk in the general population, such as cigarette smoking, obesity, lack of physical activity, and dietary factors such as high alcohol intake.[14] Underlying medical conditions that led to organ transplantation may also increase cancer risk. For example, lung recipients with cystic fibrosis and liver recipients with primary sclerosing cholangitis have an increased risk of colorectal cancer (CRC).[15]

Kidney and lung recipients usually only receive a single transplant and one native organ is left in place, therefore a chronic disease that led to organ transplantation may still affect the native organ. Among kidney recipients, about 89% of kidney cancers arise in the native kidney which may be due to kidney diseases such as acquired cystic kidney disease.[16] Similarly, the majority of lung cancers among lung recipients develop in the native lung. These are likely caused by smoking-related lung disease such as chronic pulmonary obstructive disease,[17] or other lung conditions such as idiopathic pulmonary fibrosis.

In addition, immunosuppression has a role in cancer development. The mechanism of cancer development is a combination of depressed immunosurveillance of the body to attack cancer cells, susceptibility to oncogenic viruses, and possibly direct effects of certain immunosuppressant medications.[18] Immunosurveillance occurs when small numbers of precancerous or cancerous cells accumulate in the body, which may provoke the immune system to respond and suppress the growth of these cells or eliminate them.[19] In one study, mice who lacked natural killer T cells, T and B lymphocytes, rapidly developed sarcomas after injection with a carcinogen compared to wild type mice. They also developed more spontaneous tumors than the wild-type mice.[20] Immunosurveillance also prevents the early stages of metastatic tumor development such as lymphatic invasion, vascular emboli, and perineural invasion.[18] Although decreased immunosurveillance likely plays a role in carcinogenesis in transplant recipients, it is difficult to quantify the independent effect since it is confounded with increased susceptibility to viruses and the presence of other cancer risk factors.

Viruses have been associated with various malignancies. There are two primary mechanisms by which viruses cause cancer: direct and indirect. Viruses that act directly infect a progenitor or clonal tumor cell and then the viral genome persists in the tumor cells. Alternatively, a virus may exert an indirect effect on cell development which predisposes the cell to tumor development.[21] Indirect mechanisms require other factors to impact the process of carcinogenesis. For example, a virus may weaken the immune system to where it becomes vulnerable and unable to control cancer growth. Viruses that have been associated with malignancy include Epstein-Barr virus which causes Hodgkin and non-Hodgkin lymphoma (NHL), hepatitis C and B viruses which cause liver cancer, human papillomavirus which causes ano-genital cancers and human herpesvirus 8 which causes Kaposi sarcoma.[22]

The direct effect of immunosuppressant medications and cancer risk is less clear, however the shift in regimens commonly used in clinical care may lead to different cancer risk profiles over time. For example, cyclosporine and azathioprine are both known to have cancer promoting effects.[18] Azathioprine is known to have a synergistic effect with UVA radiation in generating mutagenic oxidative DNA damage. Thus this medication may be associated with increased risk of skin cancer by inducing a chronic oxidative stress. Other medications such as tacrolimus and mycophenolate, now used by a majority of recipients, have less conclusive evidence on their association with cancer incidence.[18]

Furthermore, the duration and intensity of immunotherapy plays a role in malignancy after transplant. Thus heart and lung recipients, who generally receive very intensive immunosuppression, have a higher risk of cancer compared to kidney recipients.[1,18] Cancers that develop post-transplant may also be more aggressive and may have accelerated growth and spread.[18] Although, transplant recipients may switch to a different or less intense immunosuppression regimen following a cancer diagnosis, they will still require immunosuppression to prevent rejection of the transplanted organ.

C. Epidemiology of cancer among solid organ transplant recipients

Transplant recipients are at two-fold higher risk of developing cancer compared to the general population. From 1987 to 2008, cancer incidence among transplant recipients was 1375 per 100,000 person-years, which corresponded to an absolute excess risk of 719 per 100,000 person-years. Solid-organ transplantation has been associated with increased risk of both infection and non-infection related cancers.[1] In analyses using population-based data with recipients transplanted from 1987 to 2008, the incidence of non-Hodgkin lymphoma was elevated 7.5-fold compared to the general population, the risk of penile and vulva cancers was also elevated.[1] Also, even after excluding cancers within 6 months from transplant, the risk of liver cancer was elevated 1.5-fold among liver recipients.[23] Risks of several non-infection related cancers also were elevated including a two-fold increase in lung cancer, 4.6-fold increase in kidney cancer, 2.4-fold increase in melanoma, and a 3-fold increase in kidney cancer, among others.

As the population of transplant recipients grows, the number of recipients who develop cancer and subsequently die from cancer will also increase making this an important public health concern. Several studies conducted outside of the US have shown increased cancer mortality among recipients of kidney,[2-4] liver, heart and lung transplants.[2,5] These studies have consistently shown that risk of cancer death is 2 to 3 times more likely among transplant recipients compared to the general population. Specifically, results from a recent Canadian study of kidney, liver, heart and lung recipients showed that the mortality risk from cancer was higher than the general population (SMR: 2.84, 95% CI: 2.18, 3.07).[2] Immunosuppression may particularly increase the risk of death for cancers of infectious etiology. Indeed, this study found an elevated risk of death for cancers presumed to be of infectious etiology (SMR: 6.6, 95% CI: 5.76, 7.25). In contrast, the only previous study in the US to estimate excess cancer mortality did not find an increased risk of cancer death among kidney recipients (SMR: 0.96, 95% CI: 0.92, 1.0).[24] However, this study relied on cause of death reported to the US Renal Data System which was unknown for 41% of the transplant recipients. These deaths were classified as non-cancer deaths which leaves a substantial opportunity for misclassification bias.

D. Metrics to quantify the cancer mortality burden

Cancer mortality is an important measure of burden since it reflects the final downstream outcome in a population accounting for cancer incidence, treatment, and survival following a cancer diagnosis. Three dimensions of the cancer mortality burden are estimated in this dissertation: cancer-attributable mortality rates, which are the numbers of deaths due to cancer divided by the follow-up time in the population; the proportion of deaths due to cancer; and the LYL due to cancer. Each of these metrics provides a different description of cancer mortality in the population.

The cancer-attributable rate is a measure of absolute effect and reflects the magnitude of cancer deaths. It measures the instantaneous rate at which deaths are occurring in the population and is expressed as the number of deaths over 100,000 person-years of follow-up.[25] A limitation of rates is that they do not reflect the size of the population at risk. They are, however, useful for comparing the impact of cancer on mortality and comparing this impact among subgroups of the population with different distributions of characteristics. The rates may also be age- and sex-standardized to facilitate this comparison among subgroups with potentially different age distributions.

The proportion of deaths due to cancer among transplant recipients, also known as the population attributable fraction, PAF, is a relative measure of impact since it is a function of the risk for non-cancer causes of death. The PAF measures the extent to which cancer causes death in the whole population in relation to overall mortality.

LYL estimate the extent to which life is shortened by cancer. Conceptually, this is the difference between the observed lifetime in the population with cancer and the expected lifetime in the population had no cancer occured.[26] Thus, this quantity accounts for both the number of people who died and the age at death. The LYL are larger for cancers that cause death among children compared to cancers that lead to death among adults, since children have longer potential future lifetimes. In contrast to mortality rates, which quantify the number of deaths occurring in a population, LYL estimates are often used to evaluate the social and economic burden that result from death. Furthermore, the LYL estimates indicate which cancers are causing premature deaths.

E. Methods to quantify the cancer mortality burden

Cancer mortality burden was estimated using data from the National Cancer Institute's Transplant Cancer Match Study, which is a linkage of the US Scientific Registry of Transplant Recipients (SRTR) and 17 population-based state and regional cancer registries. All US transplant centers and organ procurement organizations submit data to SRTR for patients on the organ wait list. Baseline characteristics include demographics, patient characteristics used to determine organ allocation, medical indications for transplant, characteristics of the organ, immunosuppressant treatment. Additional data are provided at the time of transplantation. Outcome data are also collected such as graft failure and vital status of the patient.

The metrics to describe the cancer mortality burden were derived for specific subgroups of the transplant population and by cancer site. Subgroups were defined by recipient characteristics such as age, organ received, year of transplant, and time since transplant. The cancer sites examined were those common in the transplant population such as lung and bronchus, prostate, female breast, kidney, non-Hodgkin lymphoma, CRC, melanoma, and thyroid cancer. Examining mortality by cancer site facilitated an understanding of which cancers are causing the greatest mortality burden or loss of life. For example, some cancers have a high risk of death but may not result in a high rate or proportion of deaths due to that cancer if that cancer is rare in the population. Conversely, cancers that have a high incidence rate may not result in a large mortality burden if they are not particularly fatal.

The estimates by transplanted organ improve understanding of which cancers are of particular concern to each organ subgroup. This informs screening guidelines, which are likely to be most effective if they are tailored to the specific needs of organ recipients.

I also estimated mortality by time since transplant, which informs patient care as patients are surviving longer and requiring longer clinical follow-up care. These results may also shed light on the impact of the duration of immunosuppression on cancer mortality.

The most straightforward approach to estimating the cancer mortality rate would be to directly use cause of death information to classify deaths as cancer or non-cancer related. The cause of death information would be used to calculate the mortality rates for those who died from cancer and those who died of other causes. However, estimates derived using this method are susceptible to bias since cause of death information is often misclassified or unknown. The US study of mortality among kidney recipients relied on cause of death reported to the US Renal Data System.[24] Cause of death was unknown for 41% of the cases and is likely of low accuracy since these data are entered by clerks who may not have adequate clinical knowledge, training, or supervision. Cause of death information is also available from death certificates but may also be inaccurate or incomplete. Inaccuracies arise since the individual who completes the death certificate must assign a single underlying cause of death which can be difficult to do since a death may be precipitated by multiple causes and thus coding is a subjective judgement.[27] This is especially true in the population of transplant recipients who may have multiple other health conditions. Although the contributing causes of death are recorded on death certificates, the National Cancer Institute does not have access to contributing CODs for national reporting of death rates based on death certificates.

An alternative approach used in aim 1.1 that avoids using cause of death information is to estimate cancer-attributable mortality among transplant recipients using attributable risk methodology.[28] This method is based on a comparison of the all-cause mortality among those with cancer to those without cancer and assumes that the excess deaths are caused by the cancer. A similar approach was used by Simard et al.[29] to estimate the cancer-attributable mortality among people with AIDS. The calculation of the population attributable risk depended on individuals being classified with respect to the outcome (vital status) and exposure (cancer). Individuals in this study were followed for five years and so their outcome and exposure status could change over the study period. The authors divided the study period into small intervals in which the outcome and exposure were reasonably assumed to be stable, calculated the deaths due to cancer in each interval, and then combined those over the entire follow-up.

A modification of this method was developed by Engels et al. and recently applied to people with HIV infection.[30,31] They used a formula for the attributable risk that requires estimation of the proportion of deaths with a prior cancer diagnosis. Since this quantity does not fluctuate over the study period, the estimation does not need to be performed separately for different intervals and then combined to obtain the final estimate. There is a computational advantage to this approach since it calculates one estimate for the entire study period. Furthermore, it avoids choosing the length of time intervals in which the outcome and exposure are assumed to be constant.

Cause of death information collected by the transplant registry is available for all recipients with and without cancer but is based on clinical information reported to transplant centers and is likely inaccurate. Especially among individuals with several health conditions, attribution to a single cause may be difficult and the death may be misattributed. In the context of cancer, a death could be misattributed to the site of metastasis instead of the primary tumor site.[32]

Accuracy of cause of death collected by cancer registries has been evaluated. The concordance between cancer site recorded on death certificates and cancer diagnosis, was evaluated by Percy et al.[32] using deaths from 1970-71 and was more recently updated by German et al.[33] using deaths from 1993-2004. Both analyses evaluated the accuracy of death certificates by comparing the underlying cause of death on the death certificate with the cancer diagnosis. If the two sources, hospital diagnosis and death certificate, indicated the same cancer site, then there was evidence that mortality statistics based on death certificates were accurate. Although, both studies found that the overall agreement was good, both over 80%, the accuracy varied by cancer site. For example, cancer sites common for metastasis such as bone, liver, and brain were often overreported by death certificates. Furthermore, if the extent of misclassification changes over time then any time trends will be distorted.[33]

Cause of death on the death certificate may be misclassified by attributing the death to another medical condition even though the underlying cause is cancer. A study conducted by Welch and Black found that about 41% of deaths within one month of cancer-directed surgery were not attributed to cancer.[34] This estimate assumes that deaths within a short time interval after surgery as treatment for cancer should be attributed to the cancer and not as a complication from surgery. To my knowledge, no other recent studies have estimated the extent of misclassification of deaths due to cancer or its treatment that are coded as non-cancer causes. It is likely that cancer deaths are undercounted especially for deaths caused by cancer or cancer treatment that occur long after initial diagnosis such as sequalae from radiation or chemotherapy. These inherent limitations of cause of death information have led to the development of statistical techniques to estimate cause-specific mortality without relying on cause of death information.[35,36]

In contrast, cancer registries use algorithms to process causes of death from death certificates to identify a single disease-specific underlying cause of death. Howlader et al.[37] developed an algorithm that more accurately captures cancer death. The algorithm considers other causes of death along with the site of the original cancer diagnosis, tumor sequence (e.g. only one primary tumor or first of several tumors), and some comorbidities such as AIDS. I leveraged this algorithm designed to reclassify cancer deaths that may be misattributed. The cause of death information collected by cancer registries for those with incident cancers was used to estimate the cancer mortality rate in aim 1.2 and used as a comparison to the estimates from aim 1.1.

Previous work in transplant recipients has estimated life-years gained from kidney, liver, and lung transplants.[38] The methods proposed in aim 2 were an extension of this work by estimating the potential life-years that were lost among those transplant recipients who developed and died from cancer. Both approaches to estimating LYL in this dissertation rely on estimating the counterfactual survival experience. This counterfactual represents the survival experience of the transplant population assuming that death due to cancer cannot occur.

The approach used in aim 2.1 was to construct a cancer-free cohort using matching methods. The primary method in this dissertation was to match on the prognostic score. Prognostic scores, also known as disease risk scores, are used to balance the baseline risk of the outcome between cancer and non-cancer transplant recipients by summarizing the association between covariates and the potential response (death).[39-41] This method was an extension of Li et al.[42] They developed methodology that established a counterfactual cohort by matching on the prognostic score, the propensity score, or both simultaneously. They used this method to contrast the average post-transplant survival function with the average survival function that would have been observed had kidney transplantation not been available.

This method was well-suited to this application since it accommodates time-to-event data, the time-dependent nature of cancer incidence over the study period, and dependent censoring. However, the final estimate derived was the difference in survival curves among the transplanted and non-transplanted. For the approach in aim 2.1, this dissertation extended this work by estimating the difference in area between the two survival curves, which is the difference in mean life-years between the cancer and non-cancer cohorts. The resulting mean is the total LYL due to cancer averaged over the number of recipients with cancer.

The approach used in aim 2.2 was to estimate the survival curves for the overall cohort and the hypothetical cancer-free cohort using Cox proportional hazards regression models. The mean LYL due to cancer among transplant recipients can be calculated as the area between the survival curve of the overall cohort (with and without cancer) and the counterfactual cancer-free cohort. This quantity is the total LYL due to cancer averaged over the number of transplant recipients. This area is equivalent to finding a difference in the mean survival time between the two groups, usually truncated at some time point to facilitate estimation. Several methods to compare the mean survival time have been proposed. Zucker[43] developed inference based on stratified Cox models in which the effect of cancer can vary over time but the other covariate effects are constant and do not vary between cancer and cancer-free groups. Chen and Tsiatis[44] also developed methodology to compare the mean survival time but assumed separate Cox models for the cancer and cancer-free groups. Finally, Zhang and Schaubel[45] proposed using flexible Cox models with group-specific covariate estimation to compare the restricted mean lifetime and their method can also accommodate dependent censoring.

Although each of these methods compares the restricted mean lifetime of two groups, they all assume that the two groups are independent. In this dissertation, the mean survival of the overall cohort and the survival of the cancer-free cohort are not independent as they are each estimated using some of the same patients. The only difference is that the cancer-free curve is estimated by censoring at cancer incidence. Since the two groups are not independent, the variance for the LYL is calculated using a bootstrap resampling approach that accommodates the correlations.

F. Significance

Understanding the effect of cancer on overall mortality in the US transplant population is critical especially as the transplant population grows with improving overall survival. Population-based estimates and trends are necessary to assess the magnitude of the cancer burden in the transplant population and also to monitor the cancer burden over time. Mortality is used to monitor the progress against cancer, ranking which cancers are contributing to the greatest burden and also to evaluate disparities between population subgroups.[46,47] Even in the era of precision medicine, population-based data are needed to identify groups of individuals with differing risk for cancer death. Furthermore, current cancer screening guidelines for transplant recipients are limited since they are based on kidney recipients and, to a lesser extent, liver recipients.[48] The mortality estimates identify subgroups with high risk of cancer death and thus assist in the clinical care of these patients and also inform public planning by enabling more efficient and targeted cancer screening or lifestyle interventions.

In addition to providing population-based cancer-attributable mortality rate estimates that are generalizable to the US transplant population, this dissertation developed methodology that can be applied to other population-based registry data. As part of my work in the Surveillance, Epidemiology, and End Results (SEER) program, I develop and apply epidemiologic methods to overcome specific challenges presented by cancer registry data. Development of such methods is critical to estimating and ultimately communicating the most accurate and up-to-date cancer statistics to the public. The SEER program also produces reports for the public such as Fact Sheets and reports of special interest such as cancer prevalence and cost of care projections.[49] The matching or modeling methods developed in this dissertation for LYL could be added and used as part of routine reporting to describe LYL from cancer in the US general population. The estimates on the transplant population could also be highlighted as a special population to describe their unique cancer profile. Reports such as these are used by the public and clinical researchers and physicians who use these estimates for their own research or in clinical practice. Finally, these methods could be used to estimate mortality due to HIV or cardiovascular disease using registry data. Also, these methods may be used in other large cohort studies in which cause of death may not be accurately collected but all-cause mortality is available.

G. Innovation

The primary innovation of this work is the application and development of methods to overcome challenges inherent in estimating mortality burden using registry data. In particular, this dissertation applied three methodologies in a unique way. The first methodology is to calculate cancer-attributable rates and proportion of deaths using the PAF. This dissertation is the first to apply these methods to the US transplant population and overcomes inherent limitations in using cause of death collected by the transplant registry.

The second methodologic development is a novel use of the algorithm that re-classifies deaths in the cancer registry. The original purpose of this algorithm was to calculate the survival for a specific cancer in the absence of other causes of death. This is a measure of net survival and when COD is known it can be used directly to estimate the cause-specific survival among cancer patients. In a cohort of cancer patients, the denominator is comprised of newly diagnosed cancer patients. In contrast, I calculated a mortality rate in which the denominator is comprised of the entire transplant population, with and without cancer. This analysis is the first application of this algorithm other than its original purpose. The results were compared with those using the PAF methodology. The comparison helped inform the interpretation and reliability of the results from the PAF method.

The third area of methods development is the estimation of LYL due to cancer in this population. The matching method has not been applied to estimate the LYL. The modeling approach is also a novel application of establishing a counterfactual control using modeling. This is a novel approach in that I used the same cohort of recipients to estimate a cancer-free counterfactual. This approach was modified to accommodate that the same recipients will contribute to the cancer-free and overall survival experience.

The goal of developing these methods is to estimate the cancer mortality burden in the transplant population. This dissertation indeed provides estimates of three metrics to describe different dimensions of cancer mortality. These estimates fill a gap in the literature since they are not currently available and contribute to a deeper understanding of cancer mortality in this population.

3. Study design and methods

A. Data source for all aims: Transplant Cancer Match study: A linkage between US transplant registry and cancer registries

This dissertation used data previously collected by the Transplant Cancer Match (TCM) Study, which links the US Organ Procurement and Transplantation Network's Scientific Registry of Transplant Recipients (SRTR) with state and regional cancer registries. Linkage is currently complete and the calendar years of coverage vary by registry, ranging from 15 to 28 years of linked data. The TCM Study includes more than 290,000 solid organ transplant recipients with information on subsequent cancer diagnoses, which represents approximately 51% of the US transplant population during this period.

All US transplant centers and organ procurement organizations submit data to SRTR on patient characteristics, transplant and outcomes. Baseline characteristics include demographics, medical indication for transplant, and characteristics of the organ. Outcomes include recipients' vital status and organ function 6 months following transplant and each year thereafter. The SRTR also links with the Social Security Master File to ascertain death in transplant patients who are lost to follow-up.

State and regional cancer registries collect data on all incident cancers excluding non-melanoma skin cancer occurring in their catchment area. Information reported to the cancer registries includes cancer organ site, histology, and first course of cancer treatment. The TCM Study includes data from 17 US population-based cancer registries, covering the states of California (1988-2012), Colorado (1988-2009), Connecticut (1973-2009), Florida (1981-2009), Georgia (1995-2010), Hawaii (1973-2007), Illinois (1986-2013), Iowa (1973-2009), Kentucky (1995-2011), Michigan (1985-2009), New Jersey (1979-2010), New York (1976- 2010), North Carolina (1990-2010), Pennsylvania (1985-

2013), Texas (1995-2010), Utah (1973-2008), and the Seattle-Puget Sound area of Washington State (1974- 2014). The aims in this dissertation included data up to 2014.

The linkage between the SRTR and cancer registries is based on a computer-based probabilistic matching algorithm followed by a manual review of potential matches. Variables included in the matching are name, sex, date of birth, and social security number. The linked data identify at least 85% of cancers, other than non-melanoma skin cancer, in transplant recipients residing in the covered geographic areas.[50] Non-melanoma skin cancers such as squamous or basal cell carcinomas are not collected by cancer registries and therefore not included in the linkage.

To ensure the same follow-up window for cancer identification for all transplant recipients, transplant recipients were included in the TCM Study if they resided in a contributing cancer registry catchment area and received their transplant after the cancer registry start date. The registry start date is defined as the year that the cancer registry achieved complete reporting (at least 95% of cases). Finally, recipients were included in the study cohort at the time of their first transplant. Any incident cancer or death was attributed to that first transplant regardless of subsequent transplants.

B. Aim 1: Estimate proportion of deaths due to cancer and the cancer-attributable mortality rates among transplant recipients using two statistical approaches

One that does not use cause of death (COD) information and one that uses COD from the cancer registries.

Since COD information can be unreliable, estimates derived using COD are likely to be biased. The method used in aim 1.1 produces estimates of the fraction of deaths attributed to cancer in the transplant population and avoids using COD. A limitation of this method is that the estimates may be biased if there is unmeasured confounding in the relationship between cancer and death. The method used in aim 1.2 utilizes COD information. However, it only uses COD information from the linked cancer registries, incorporating information on prior incident cancer diagnoses, which is more likely to be accurate. This method is not susceptible to bias caused by unmeasured confounding but is limited by the accuracy of COD information. Since these methods have different limitations, they are complementary and were compared in aim 1.3.

1) Aim 1.1: Estimate proportion of deaths associated with cancer and the cancer-attributable mortality rate using population attributable fraction methodology

Study Design & Participants

All adult transplant recipients who received their first solid organ transplant during a period of cancer registry coverage and did not have a prior cancer other than non-melanoma skin cancer were included. Prior cancers were defined as those diagnosed before transplant or within 90 days post-transplant since these are likely to have developed before transplant.

Measures

The primary outcome was death from any cause. Month and year of death were used from the SRTR and cancer registries. This information has been reconciled and a transplant recipient is recorded as dead if either source has a record of death with the earlier death date used.

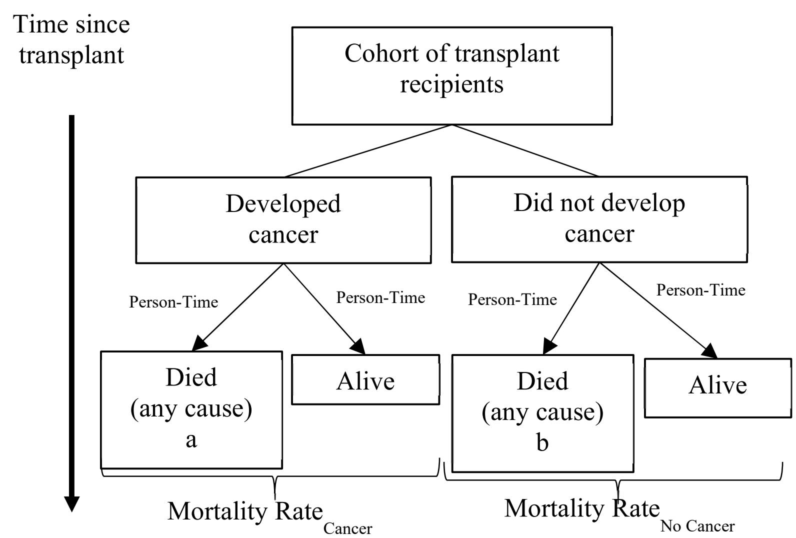
Invasive primary cancers were identified from the linked cancer registries and coded according to the International Classification for Diseases for Oncology, version 3 and classified using a modified version of the Surveillance, Epidemiology and End Results

(SEER) program site recode.[51] If a recipient had multiple primary cancers, only the first primary was included. Excluding non-melanoma skin cancer, most transplant recipients with cancer had only one primary cancer (>90%).

Data on confounders was collected by the SRTR. They include transplanted organ and year of transplant. Demographic characteristics such as sex, race/ethnicity and age were included.

Analysis

The population attributable fraction (PAF) associated with cancer is the fraction of deaths attributable to cancer among all transplant recipients. Specifically, $PAF = p_d \times AR_{Cancer}$ where p_d is the number of deaths among those with a prior cancer diagnosis divided by all deaths in the cohort and AR_{Cancer} is the proportion of deaths attributable to cancer among people with cancer. I estimated AR_{Cancer} with (RR-1)/RR where RR is the relative risk of all-cause mortality associated with cancer compared to those without cancer.[28] This avoids using the COD. These quantities are depicted conceptually in Figure 1. The RR was estimated using an adjusted Cox proportional hazards regression model with cancer as a time-dependent variable to ensure that person-years without cancer among those who subsequently develop cancer were counted as cancer-free time. The at-risk time for each person started 90 days post-transplant. The follow-up period ended at death, loss to follow-up by the SRTR or end of cancer registry coverage, whichever occurred first. Death from any cause was counted as an event, while transplant recipients lost to follow-up or alive at the end of cancer registry coverage were censored. The follow-up time started 90 days after the first transplant and ended at death or a censoring event, regardless of any subsequent transplants or non-functioning period for kidney recipients. Time since transplant was used as the time scale with 90 days as the time origin since this is the time when recipients become at risk for a post-transplant cancer and subsequent death. I considered only first cancers in the PAF calculations.



Attributable Risk$_{\text{Cancer}}$

$$= \frac{\text{Mort Rate}_{\text{Cancer}} - \text{Mort Rate}_{\text{No Cancer}}}{\text{Mort Rate}_{\text{Cancer}}}$$

= (RR-1)/RR

Percentage of deaths with a cancer diagnosis

p_d= a/(a+b) where a = number of transplant recipients with cancer who died, b=number of transplant recipients without cancer who died

Figure 1: Population attributable fraction calculation

The models were adjusted for covariates such as sex, race/ethnicity, age at transplant, calendar year of transplant, time since transplant, and organ. Since age is strongly associated with both risk of cancer and death, the functional form was thoroughly investigated (Appendix 1).

If the proportional hazards assumption was violated then the approach of Simard et al.[29] could have been used. However, the proportional hazards assumption of the models was checked by including an interaction term between cancer and follow-up time in the model and this assumption help (Appendix 1.B). The other covariates in the model also were checked (Appendix 1.C). However, non-proportionality of the hazards for other covariates should not strongly impact the RR estimate for cancer. Another assumption of the Cox model is that censoring is noninformative conditional on the covariates in the Cox model. That is, that the censoring mechanisms are not related to the probability of death. The linearity assumption of the functional form was evaluated by plotting the Shoenfeld residuals.[52]

Primary analysis included all transplant recipients and all cancers combined. In secondary analyses, I also estimated the PAF for various subgroups of the transplant population, such as by age group, sex, time since transplant, and organ received. A separate model was estimated for each subgroup with appropriate model fitting and diagnostics for each. I also estimated the PAF for individual cancer sites that are common in the transplant population. For example, the PAF was estimated for lung and bronchus, prostate, female breast, kidney, non-Hodgkin lymphoma (NHL), CRC, melanoma, and thyroid cancer. The PAF also was estimated for all other cancers not listed above as a group.

PAF estimates were multiplied by the overall mortality rate to obtain the mortality rate among transplant recipients attributed to cancer (Equation 1). I also tested for differences between subgroups such as differences by age, cancer site and time since transplant. Methodology to estimate the variance and confidence intervals has been derived by Engels et al.[30]

$$\frac{Proportion\ of\ deaths\ due\ to\ cancer \times Total\ number\ of\ deaths}{Person - Years_{Transplant\ Coho}}$$

Equation 1: Cancer-attributable mortality rate as computed in aim 1.1

Several sensitivity analyses were conducted. The first was to evaluate the exclusion of cancer occurring within 90 days post-transplant. Although the 90 day window is likely conservative since previous literature on cancer mortality has either not excluded any cancer cases post-transplant[3,4,24] or used a 30 day window,[2] I evaluated alternative exclusion windows up to 1 year on the estimation of the relative risk (Appendix 1.D). Next, I evaluated how much the PAF estimates may be affected by unmeasured confounding. While there may be some residual confounding in the estimates of the relative risk, I did not anticipate a large impact on the PAF. The relative risk between cancer and death is large for many cancer sites and so mathematically the factor (RR-1)/RR is likely close to 1 and PAF$\approx p_d$. It is unlikely that unmeasured confounding would bias the relative risk enough to substantially change the PAF estimate. However, I conducted a sensitivity analysis for cancer sites where the estimated relative risk was lower than expected, or for cancer sites for which the relative risk of death was not very strong (e.g. breast or prostate cancer).

Missing data is a potential problem in any epidemiologic study. There was some missing data for the covariates used in the models. This was expected to be very low amount of missing data since the registries have fairly complete collection of these data. Recipients with missing data were excluded from the modeling. If the amount of missing data was causing 10% or more of recipients to be excluded, then multiple imputation was used.[53]

Precision Estimate

The cohort contains data on about 290,000 transplant recipients with about 11,000 having a linked cancer diagnosis. Approximately 62,600 deaths were observed by 2016. It is difficult to estimate the precision of the PAF before analyzing the data since the PAF relies on the hazard ratio, the proportion of deaths with a cancer diagnosis and the correlation between them. Two approaches were used to approximate the precision. The first approach was to use preliminary data from a cohort of kidney transplant recipients to estimate the PAF and its Wald confidence interval.[54] As shown in Table 1 the confidence intervals are very narrow. This interval estimate is very crude as it does not account for the correlation between the hazard ratio and the proportion of deaths with a cancer diagnosis, and therefore, the confidence intervals are likely too narrow. Alternatively, results are available from a study of people with HIV, another immunosuppressed population. Here the PAF and the confidence intervals were estimated using correct methodology (Table 1). The PAFs from both cohorts are very similar. The confidence intervals for the HIV cohort are wider but more accurate since they account for the underlying correlation in the PAF estimation.

The estimates obtained in aim 1 have at least the precision of the HIV estimates and were expected to be more precise since the cohort of transplant recipients is larger. For example, there are 2775 people with NHL, 2824 people with lung cancer, and 1613 people with liver cancer. Thus, the estimates obtained using the cohort of transplant recipients are even more precise.

	Cohort of Kidney Recipients Method of Jewell (2004)[18]			Cohort of people with HIV Method of Engels et al.[17]		
Cancer Site	Total deaths	PAF	95% CI*	Total deaths	PAF	95% CI
All	4421	11.48	11.40, 11.56	1069	9.8	9.1, 10.5
NHL	583	1.69	1.67, 1.70	207	2.0	1.7, 2.3
Lung	738	2.56	2.54, 2.57	222	2.3	2.0, 2.6
Liver	74	0.26	0.26, 0.27	80	0.9	0.7, 1.1

*The 95% confidence intervals were computed using the EpiR[55] package in R

Table 1: Estimated population attributable fraction and its 95% confidence interval for two cohorts

Strengths

A strength of this method is that the PAF is estimated using the relative risk for overall mortality associated with cancer in transplant recipients to estimate cancer and non-cancer specific mortality rates. This is an innovative application that avoids using potentially misclassified COD from death certificates and uses estimates of the relative risk of death from cancer from a statistical model that is adjusted for confounders. The advantage of this method is that it depends only on the proportion of deaths among transplant recipients with a prior history of cancer. The proportion of deaths among transplant recipients with a prior history of cancer was evaluated for changes over time.

This cohort has a large sample size that enables estimation of the PAF by transplanted organ and cancer site. Since cancer risks differ by transplanted organ and mortality rates differ by cancer site, stratified analyses may be more accurate. Finally, because of excellent coverage in the TCM Study, PAF estimates from this cohort data are generalizable to the US transplant population.

Limitations

Transplant recipients may have increased cancer detection due to heightened medical surveillance. If this is the case, the estimated relative risk of death among transplant recipients with cancer compared to those without cancer may be underestimated since more indolent and early stage cancers will be diagnosed. However, the proportion of recipients who died of any cause, p_d, among those with an incident cancer will be higher. Since the RR would be biased toward the null but p_d would be biased away from the null, the PAF may be biased in either direction. These biases are not expected to be large. Furthermore, mortality estimates were estimated by cancer site, which creates groups of recipients that are more homogeneous with respect to other comorbidities and medical surveillance. Finally, the estimates based on the relative risk and p_d were compared to estimates using COD. Large differences may highlight where these assumptions may not be met.

PAF methodology generally assumes that there is a causal relationship between cancer and mortality with no unmeasured confounding. This assumption may not hold in this analysis where data on potential confounders including other comorbid conditions and risk factors such as smoking were not available. For example, if another medical condition is associated with both a specific cancer and death, the PAF may be biased away from the null since deaths will appear to be associated with cancer when, in fact, they are due to this other condition. The result of this bias is expected to be small since the relative risk between cancer and death is large for many cancer sites, and mathematically the factor (RR-1)/RR is likely close to 1 and PAF$\approx p_d$. To evaluate the impact of unmeasured confounding, a sensitivity analysis was performed as described above. This provides a range of plausible values for the PAF estimates.

Since cancer registries do not collect information on non-melanoma skin cancers, these were not included in the analysis. In particular, squamous cell carcinoma has an elevated incidence in transplant recipients[56,57] While these skin cancers are a contributor to morbidity among transplant recipients, these cancers are not often fatal, and therefore, do not greatly affect mortality estimates.

Finally, some cancers may be missed. This may happen if a cancer is not reported to the cancer registry, if the linkage does not match the cancer to the transplant recipient, or if a transplant recipient out-migrated from the catchment area. However, out-migration is uncommon, with 94% of recipients living in the same state 10 years after transplant.[1]

2) Aim 1.2: Estimate the proportion of deaths due to cancer and the cancer-attributable mortality rate using COD from the cancer registries

As discussed above, COD information from the transplant registry is likely to be inaccurate or incomplete and information on incident cancer is not reliable. However, an algorithm has been developed using COD information from cancer registries that corrects for inaccuracies in the death certificate coding among people with cancer.[37] This relies on the presence of the incident cancer diagnoses in the cancer registries. Specifically, if a cancer diagnosis is documented prior to a COD of cancer, then the likelihood that the person died of cancer is high. In this aim, I leveraged the higher accuracy of the COD available from the cancer registries compared to COD collected by the transplant registry.

Study Design & Participants

The participants included were the same as in aim 1.1.

Measures

COD is obtained by the cancer registries by a linkage with the National Death Index. Invasive cancers were identified as described in aim 1.1. The primary outcome was death from cancer as recorded by the cancer registry. The algorithm used by cancer registries to reclassify cancer deaths is different for those diagnosed with a single primary cancer and those with more than one primary cancer.[37] In patients with only one cancer diagnosis, any cancer death is attributed to the same site as the incident cancer. Also, for some cancer sites, deaths coded as HIV related and deaths due to other diseases that were related to the cancer are reclassified to a death due to the primary cancer site. Examples of other noncancer diseases include a death due to liver cirrhosis in a patient with liver cancer, certain ulcer conditions in a patient with CRC, stomach or esophageal cancer, and certain diseases of the breast in breast cancer patients.

For patients who had more than one cancer, the algorithm reclassifies cause of death for the first cancer. The rules include classifying cancer deaths to the site of the first primary cancer if the underlying cause of death was attributed to the same organ system or was attributed to multiple cancers with an unknown primary. The same rules for noncancer causes of death are applied for those with more than one cancer and a single primary. Finally, deaths from all other malignant cancers were not classified as cancer deaths since they presumably are correctly attributed to subsequent incidence cancer sites.

Analysis

The number of transplant recipients who had a linked cancer diagnosis, died, and had a COD listed as cancer by the cancer registry was divided by the person-time at risk to derive cancer-attributable mortality rates (Equation 2). I also estimated the PAF by dividing the number of deaths with a COD of cancer by the total number of deaths. These estimates were produced for the same subgroups as in aim 1.1.

$$\frac{Number\ of\ cancer\ deaths}{Person - Years_{Transplant\ Cohort}}$$

Equation 2: Cancer-attributable mortality rate as computed in aim 1.2

Sample Size

The same sample was used as described in aim 1.1.

Strengths

The strength of this method is that it uses COD collected by the cancer registries which is more accurate than the CODs in the transplant registry. This method is also not affected by confounding of the relationship between cancer and death. Finally, missing data on recipient characteristics does not affect the results since these covariates are not needed for modeling.

Limitations

The approach assumes that transplant recipients without a linked cancer diagnosis did not die from cancer. As in aim 1.1, some cancers may be missed. This would result in an underestimate of the cancer-attributable mortality rate. Finally, there may be errors in the underlying COD entered by the physician. The direction of the bias would be an underestimation since recipients with cancer may be miscoded to noncancer deaths. The algorithm mitigates these errors, but some may be present especially for deaths caused by late effects of cancer treatment.

3) Aim 1.3: Compare estimates from aim 1.1 and 1.2

The cancer-attributable mortality rates from aim 1.1 and aim 1.2 were compared overall and by major subgroups defined by transplanted organ, other recipient characteristics, and for the most common cancer sites. The rates were compared by assessing if their 95% confidence intervals are overlapping. A formal statistical test was not performed since the two comparison groups are not independent (i.e. they are estimates using the same data produced with different methodology). Furthermore, the objective is not to determine statistically significant differences in estimates derived using the two methods, but rather, to evaluate the sensitivity of the estimates to the method used since the estimates should be similar. I discuss any potential differences. For example, if they are not similar, the PAF in aim 1.1 could be wrong due to unmeasured confounding. Alternatively, the cancer-specific deaths obtained from cancer registries used in aim 1.2 could be incorrect. An incorrect COD may be due to an error in the true underlying COD entered by the physician or a misclassification by the algorithm. I explored possible reasons for any discrepancies. For example, large differences for cancer sites with risk factors that are not measured in the transplant registry may be indicative of a PAF estimate with residual confounding.

C. Aim 2: Estimate mean LYL due to cancer among transplant recipients

Two methods were used to estimate LYL, which differ on the estimation of survival had cancer not occurred in the cohort. The mean LYL due to cancer among those who develop cancer also was estimated.

Transplant recipients with cancer are assumed to have shorter lifetimes than they would have had without cancer. The LYL due to cancer is the difference between the actual lifetime with cancer and what that recipient would have experienced without cancer. Thus, calculating the LYL due to cancer requires estimation of the hypothetical lifetime those with cancer would have lived had they not developed cancer. This is referred to as the cancer-free counterfactual experience of those with cancer. In this aim, the mean LYL due to cancer was estimated using two methods. The first method, described in aim 2.1, matches each recipient with cancer to those without cancer. The matched cancer-free recipients were used to estimate the lifetime that would have been experienced by the recipients had they not developed cancer. The second method, described in aim 2.2, used survival modeling to estimate the counterfactual cancer-free experience. These methods are based on different assumptions and the resulting mean LYL estimates were compared in aim 2.3.

1) **Aim 2.1: Estimate mean LYL due to cancer using matching methods**

Each person with cancer was matched to a person without cancer to provide a counterfactual estimate of the life-years lived without cancer.

Study design & participants

The study cohort was the same as described in aim 1.

Measures

The primary outcome is death from any cause. Invasive primary cancers and confounders were measured as described in aim 1.1.

Analysis

The objective is to estimate the survival function that represents the hypothetical cancer-free experience. For each transplant recipient with cancer, similar recipients without cancer were matched to serve as the counterfactual controls. The controls are intended to reflect what the survival time would have been had the cancer patient not developed cancer. The mean LYL due to cancer is the difference in the areas under the hypothetical cancer-free curve and the survival function estimated using the subset of recipients who develop cancer.

Prognostic score

Non-cancer controls were matched to cancer cases using a prognostic score. Prognostic scores, also known as disease risk scores, are used to balance the baseline risk of the outcome between cancer and non-cancer patients by summarizing the association between covariates and the potential response.[39-41] Controls were matched to each cancer patient using these scores.[42,58,59] Matching on the prognostic score creates a set of cases and controls with the same baseline risk of death (i.e. unconfounded by the set of variables that was used in estimating the score). This is achieved by eliminating the association between the confounders and the outcome (death).

The prognostic scores were estimated based on a Cox regression model using the full cohort of transplant recipients with events defined as death from any cause. Censoring occurred at the time of cancer incidence, loss to follow-up, or at 10 years post-transplant. Since censoring occurred at the time of cancer incidence, this model reflects the pre-cancer experience of the entire cohort. The prognostic score models were estimated separately by organ transplanted since the association between the covariates and death vary by organ. As in aim 1.1, the time scale used was time since transplant. The model included all predictors of pre-cancer death and potential confounders of the relationship between cancer and risk of death such as age, sex, calendar year of transplant. Although, borderline significant variables remain in the model since removal of many variables that have small amounts of confounding may accumulate and produce less optimal scores.[41] For each patient, the prognostic score was computed using their individual covariate profile. Given by, $\widehat{\varphi_l} = \hat{\beta} X_l$ where $\hat{\beta}$ is the vector of estimated log-relative risk coefficients from the prognostic Cox model and X_l is the observed covariate vector for subject *l*.

Matching procedure

As in Li et al., patient *k* was selected as a match to patient *l*, so that $\left|\log \hat{\psi}_{l:k}\right|$ is minimized where $\hat{\psi}_{l:k} = \exp\,(\hat{\beta}(X_l - X_k))$. Eligible matches had the same organ transplanted as the cancer case, were alive at the time of cancer incidence, not yet had developed cancer, and received the transplant in the same calendar year. A control was eligible even if they developed cancer at a later time, as long as they were cancer-free at the time from transplant of the case's diagnosis. The matching algorithm used matching-with-replacement in that a subject who has not yet developed cancer may be selected as a match to multiple cancer subjects. Controls were identified for each cancer patient using nearest neighbor matching within a specified caliper distance.[58,59] The matching ensured that the covariate distribution was balanced with respect to risk of non-cancer death between cancer patients and the counterfactual controls. Up to 5 eligible matches within the caliper distance were selected. If a cancer case did not have a match within the caliper distance, then that case was excluded from the analysis. In addition, the matching allowed for the same control to be selected to more than one cancer case. Also, a subsequent cancer case may be selected as a control before they developed cancer.

The optimal caliper width for matching on the propensity score is well-studied.[60,61] In general, matching on a fraction of the standard deviation of the logit of the propensity score removes a majority of the bias in the resulting treatment estimate. However, only one study has examined caliper width when matching on the prognostic score. A simulation study found that matching on the natural-scale may be too large for matching on the prognostic score.[58] Similar to the literature on propensity scores, calipers based on fractions of the standard deviation of the logit of the prognostic score performed well in removing bias from the final estimates. However, this simulation was based on a logistic regression to estimate the prognostic score as opposed to a survival outcome estimated using Cox regression.

Several caliper widths were assessed to balance finding a match for every case with using matches that were sufficiently similar to the case. The matching was evaluated based on the marginal distribution of covariates between the cancer and non-cancer groups. If there were large imbalances then additional stratification was needed (e.g. also stratify by sex). Imbalance was determined by the p-value but also differences that are clinically or substantively meaningful. Since the dataset is large, even small differences may result in a statistically significant p-value.

LYL estimation

The mean LYL is the difference between the areas under the survival curves up to the truncation time. This is equivalent to the difference in the restricted mean lifetime between the non-cancer and cancer groups. The interpretation of the restricted mean lifetime at, say, 10 years, is the average number of years lived out of the next 10. The censoring time was chosen so that the most amount of follow-up time was retained but there were sufficient events in all parts of the survival curve. I expected this point to be at about 15 years since transplant.

Survival curves were estimated using the non-parametric Kaplan-Meier approach. The mean LYL due to cancer among those who develop cancer was estimated by subtracting the area under the survival for the cancer cohort from the area under the survival curve of the hypothetical cancer-free cohort. This can be written as:

$$\overline{LYL}_{case\_only} = \overline{LYL}_{controls} - \overline{LYL}_{cancer}$$

where $\overline{LYL}_{controls}$ is the area under the hypothetical cancer-free curve estimated using data from the matched controls and $\overline{LYL}_{cancer}$ is the area under the curve estimated using data from recipients who developed cancer. The variance of the LYL was estimated using the bootstrap.

In addition to the mean LYL due to cancer, I also computed the fraction of LYL due to cancer in the transplant population by dividing the total LYL (i.e., $\overline{LYL}_{case\_only}$ x number of cancer cases) by the total life-years lived in the transplant population. Finally, the mean LYL due to cancer averaged over the number of transplants ($\overline{LYL}_{tx\_pop}$) was also computed by dividing the total LYL by the number of transplants. Both the mean LYL and the fraction of LYL were estimated by transplanted organ, cancer type, age group, and sex.

Sensitivity analyses

The following sensitivity analyses were conducted:

1) Caliper width was investigated as described above (Appendix 2.C).
2) Other scores for matching were investigated.

   The matching was performed to select controls that are very similar to the cases. Each method approaches this differently. The prognostic score creates a control set that has the same baseline risk of death as the cases, whereas the propensity score balances the covariate profile between the cases and controls. The prognostic score was chosen since it is better for rare exposures and it is known to create better overlap between cases and controls. However, I evaluated using the propensity score for matching (see sensitivity analyses in Chapter 7).

Strengths

The strength of this method is that it uses controls to serve as the counterfactual. The matching based on a score is a convenient approach to handling many covariates.

Limitations

Unmeasured confounding may bias these results. For example, confounding due to other medical conditions or risk factors may cause an overestimate of life-years lost if cancer is associated with some other fatal medical condition. That condition may lead to the loss of life years, but it will appear that it is due to cancer. This method also relies on Cox models and thus the assumption is that these models are correctly specified. Finally, this method as described cannot accommodate time-dependent confounders. Estimation of the prognostic score using the same data to which it is applied can lead to an inflated Type 1 error for the exposure (i.e. cancer).[39,40] However, this occurs when the correlation between cancer and the confounders is very high. This is unlikely to be the case for these data. Finally, since the time since transplant was truncated, the maximum LYL due to cancer for any recipient is reached at the truncation time. This results in an underestimate of the LYL due to cancer and may result in a more severe underestimate for children.

2) **Aim 2.2: Estimate the mean LYL using modeling. Cox models were used to estimate the counterfactual, which is survival had no cancer occurred among transplant recipients.**

Study design & participants

The study cohort was the same as described in aim 1.

Measures

The primary outcome is death from any cause. Invasive cancers and confounders were measured as described in aim 1.1.

Analysis

The mean LYL due to cancer in the transplant population was estimated by comparing the survival curve of the hypothetical cancer-free cohort and the survival of the overall cohort. The mean LYL due to cancer among those who develop cancer was estimated by comparing the cancer-free survival experience to the survival function for those who develop cancer. Each of these curves was estimated using Cox regression.

The life-years lived from transplant to time τ can be estimated as the area under the survival curve up to this time point. This is mathematically equivalent to the restricted mean survival time up to time τ, $\mu = \int_0^\tau S(t)dt$. The life-years lost due to cancer among transplant recipients up to time τ is the difference in the area under the cancer-free survival curve and the overall survival curve, shown graphically in Figure 2.

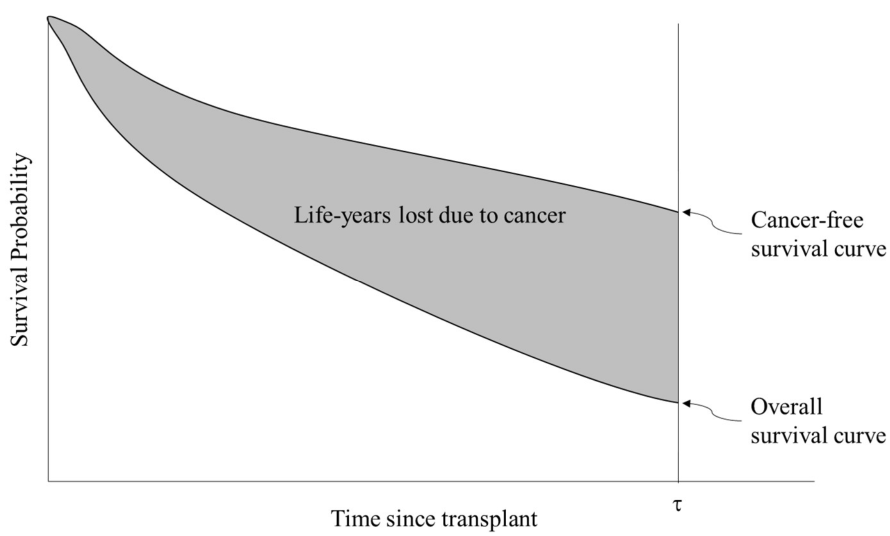


Figure 2: Life-years lost due to cancer among transplant recipients

Overall survival in the cohort was estimated using a Cox regression model adjusted for sex, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian/Pacific Islander, other), age at transplant (modeled piecewise with knots at age 5 and 19), transplant year (1995-99, 2000-04, 2005-08, 2009-14), body mass index (underweight, normal, overweight, obese), diabetes (yes, no, unknown), registry, organ-specific reasons for transplant. Events were defined as death from any cause and censoring occurred at loss to follow-up or 10 years post-transplant. The time scale was time since transplant. Incident cancer was not explicitly part of the model since the purpose is to estimate the overall survival of the cohort averaged over cancer incidence. The model can be written: $S_1(t) = exp\{-\Lambda_{01}(t)e^{\beta X}\}$ where $\Lambda_{01}(t)$ is the cumulative baseline hazard function. The baseline hazard was estimated using the Breslow method. This represents the survivor function for an individual whose covariate values are all 0. If there are many ties in the event times in the data, other methods such as the Efron method may be used instead.

A second model estimated the survival had no cancer occurred. This was estimated using a Cox model adjusted for covariates with events defined as death from any cause. In this model, censoring occurred at incident cancer, loss to follow-up, or 15 years post-transplant. The model was: $S_0(t) = exp\{-\Lambda_{00}(t)e^{\alpha X}\}$. The coefficients from the Cox models were estimated separately for the overall and cancer-free survival curves. The baseline hazard was estimated as described above.

The mean LYL due to cancer was calculated from average survival curves. Here, the average survival curve is one that is adjusted for covariates and represents the average survival experience over all subjects in the cohort. To estimate the average survival curve in each group there are several methods:

1. Estimate the survival curve for each individual and then average curves. For example, $\hat{\beta}x$ and $\hat{\alpha}x$ are estimated by using each individual's covariate vector. Then the survival functions are averaged at each event time to produce the average overall survival curve and the average curve under the cancer-free assumption. Event times that occur in either the overall survival or cancer-free assumption are used to estimate the survival functions.
2. Estimate the survival curve using the average covariate values. For example, $\hat{\beta}\bar{x}$ and $\hat{\alpha}\bar{x}$ are estimated where $\bar{x}$ is the vector of sample means for each covariate in the sample at baseline.
3. Choose a representative profile to estimate the survival curve. For example, estimate the survival curve for white men age 60.

I used option 1 since it represents the average curve in the sample. Once the average survival curves are estimated then they can be used to calculate any summary statistics of interest.

The primary outcome of this aim is the estimate the mean life-years lost due to cancer among transplant recipients. Similar to aim 2.1, these estimates are the restricted mean lifetime up to the truncation time, say 15 years. Let $\overline{LY}_0$ be the mean life-years lived in the cancer-free cohort and $\overline{LY}_1$ be the mean life-years lived in the overall cohort. Each of these quantities is estimated by the area under the average survival curve. The mean life-years lost due to cancer is the difference, i.e. $\overline{LYL}_{tx\_pop} = \overline{LY}_0 - \overline{LY}_1$. The area under the curve is truncated since the tails of the survival curves do not reach zero and would require estimation. To avoid estimation of the tails, a standard technique is to truncate the follow-up time. Since the objective is to summarize as much of the survival experience post-transplant as possible, time τ was chosen to be 15 years. The time point needs to be chosen so that it is smaller than the latest observed event time but long enough to capture as much of the survival experience as possible. If the data do not support $\tau = 15$ then it can be reduced until the area under the curve can be estimated.

The variance of $\overline{LYL}_{tx\_pop}$ was estimated using a bootstrap. The bootstrap was used since the same subjects were used to estimate each model and, therefore, the variance must account for the correlation between $\overline{LY}_0$ and $\overline{LY}_1$. The bootstrap procedure was:

1. Draw a sample from the cohort with replacement to create a bootstrap cohort of the same size as the original cohort.
2. Estimate the average survival curves for each group.
3. Estimate the area under the curves and their difference and obtain $\overline{LYL}_{tx\_pop,b}$
4. Repeat 1000 times.
5. The 95% confidence interval of the mean LYL estimate is given by the 5th and 95th percentiles of the bootstrap distribution.

The total LYL was computed by $\overline{LYL}_{tx\_pop}$ x number of transplants. From this, $\overline{LYL}_{case\_only}$ can be computed by dividing the total LYL by the number of recipients who developed cancer. The mean LYL was estimated by transplanted organ, cancer type, age group, sex and year of transplant.

Strengths

A strength of this method is that it uses the complete observed cohort of transplant recipients to estimate the survival curves. Also, the final survival curves, comparing the cancer group and the non-cancer controls, do not assume any specific functional form in modeling the effect of cancer on death. Also, I obtained hazard ratio estimates for the other confounders in the model, which enabled examination of the effect of these confounders on survival in the presence of cancer. Finally, this method is more statistically powerful than the approach in aim 2.1 since it relies on parametric modeling.

If the survival curves cross, then other measures that are time-point dependent may be conflicting since the LYL is a summary measure across the entire follow-up period.

Once the average survival curves are estimated, any other measures can be calculated. Some examples include:

1. A. Difference in t-year survival rates. These could be used to evaluate short and long-term LYL from cancer (e.g. difference in survival 12 months post-transplant and difference 5 years post-transplant).

   B. Difference in survival curves over time. Shows how the survival changes over time (e.g. may see crossing of the survival curves).
2. Ratio of percentiles of the survival function.
3. Estimated residual lifetime: Probability of survival after age a.

Limitations

The limitation of unmeasured confounding is the same as in aim 2.1. One assumption is that the survival after cancer incidence is the same as the cohort without cancer.

3) Aim 2.3: Compare LYL estimates from both methods

The results for the mean LYL are expected to be similar from both methods. If the results are not similar, then that may indicate poor matching. The modeling approach is less sensitive to limited data within a specific covariate profile but relies more on correct specification of the models and their functional forms.

4. Cancer-attributable mortality among solid organ transplant recipients in the United States, 1987 through 2014[1]

A. Abstract

Background: Solid organ transplant recipients have an elevated risk of cancer. Quantifying deaths attributable to cancer can inform priorities to reduce cancer burden.

Methods: Linked transplant and cancer registry data were used to identify incident cancers and deaths among solid organ transplant recipients in the United States (1987-2014). Population-attributable fractions (PAFs) of deaths due to cancer and corresponding cancer-attributable mortality rates were estimated using Cox models.

Results: Among 221,962 transplant recipients, 15,012 developed cancer. Thirteen percent of deaths (PAF=13.2%) were attributable to cancer, corresponding to a cancer-attributable mortality rate of 516 per 100,000 person-years. Lung cancer was the largest contributor to mortality (PAF=3.1%), followed by non-Hodgkin lymphoma (NHL, PAF=1.9%), colorectal cancer (PAF=0.7%), and kidney cancer (PAF=0.5%). Cancer-attributable mortality rates increased with age at transplantation, reaching 1229 per 100,000 person-years among recipients 65+ years old. NHL was the largest contributor among children (PAF=4.1%) and lung cancer among 50+ year-olds (PAFs=3.7-4.3%). Heart recipients had the highest PAF (16.4%), but lung recipients had the highest cancer-attributable mortality rate (1241 per 100,000 person-years). Overall, mortality attributable to cancer increased steadily with longer time since transplant, reaching 15.7% of deaths (810 per 100,000 person-years) 10+ years post-transplant. Comparison of cancer-

[1] Noone AM, Pfeiffer RM, Dorgan JF et al. (2019) 125:15. *Cancer*.

attributable mortality rates with specified causes of death indicated that some deaths recorded as other causes might instead be caused by cancer or its treatment.

Conclusions: Cancer is a substantial cause of mortality among solid organ transplant recipients, with major contributions from lung cancer and NHL. Cancer-attributable mortality increases with age and time since transplant, so cancer deaths will become an increasing burden as recipients live longer.

B. Introduction

Solid organ transplant recipients have increased cancer risk compared with the general population.[1] Excess risk varies by cancer type and is largely due to immunosuppression from medications used to prevent rejection, underlying medical comorbidities, and end-stage organ disease. Recipients have elevated risk for both virus-related cancers (e.g., non-Hodgkin lymphoma [NHL], caused by Epstein-Barr virus [EBV]) and virus-unrelated cancers (e.g., lung and kidney cancers).[1] Moreover, the transplant population in the United States has doubled since the late 1980s due to an increasing number of transplants and improved survival.[12] Understanding long-term health risks, including cancer, is critical.

Cancer mortality is an important measure of burden since it reflects the final downstream outcome accounting for cancer incidence, treatment, and survival following cancer diagnosis. For many cancers, patients with a previous transplant have an increased risk of dying from cancer compared to those without a transplant.[62,63] Immunosuppression associated with transplantation may impair control of the primary tumor or metastases. Studies conducted outside of the US have shown that the population of transplant recipients overall has a 2-3-fold elevated mortality from cancer compared to the general population.[2-5]

Cancer mortality is usually estimated using cause of death (COD) information. However, such estimates depend on accurate classification and reporting by physicians. Determining single underlying CODs can be difficult since death may be precipitated by multiple causes and determination of the underlying COD is often subjective.[27] This may be especially true for deaths in transplant recipients, who have multiple health conditions.

An alternative approach to estimating cancer-related mortality is to compare the overall mortality of individuals with and without cancer, attributing excess deaths in the former group to cancer. This statistical calculation leads to a population-attributable fraction (PAF), which describes the proportion of all deaths attributable to cancer. The PAF depends only on ascertainment of incident cancer diagnoses and overall mortality, thus avoiding the need for COD. Further, the cancer-attributable mortality rate can be derived from the PAF.

Population-based estimates can help quantify the cancer burden in the transplant population and identify specific cancers and subgroups that might benefit from intervention. In this study, we estimated the fraction of deaths attributed to cancer and corresponding cancer-attributable mortality rates among US transplant recipients.

C. Methods

The Transplant Cancer Match (TCM) Study links the US Scientific Registry of Transplant Recipients (SRTR) with state and regional cancer registries.[1] SRTR data include recipient demographic and transplant characteristics, vital status and organ function. COD is based on clinical information reported by transplant centers and updated over time, so recipients may have multiple CODs. Seventeen cancer registries provided data on incident cancers (excluding non-melanoma skin cancer) covering approximately half of the US transplant population (see Table 2 footnote). The TCM Study was approved by institutional review boards at the National Cancer Institute and participating cancer registries.

Invasive first cancers in transplant recipients were identified from the linked cancer registries and classified as previously described.[51] The study cohort comprised recipients who resided in areas covered by participating cancer registries at the time of transplant, received their first transplant during a period of cancer registry coverage, and did not previously have cancer recorded in the cancer registry. Prior cancers were those diagnosed before transplant or within 90 days post-transplant (since those likely developed pre-transplantation). Follow-up thus started 90 days post-transplant, and ended at the earliest of death, loss to follow-up by the SRTR, or end of cancer registry coverage. Since we sought to quantify the impact of cancer on mortality once a transplant is performed, we did not censor follow-up if the transplanted organ failed or the person received a subsequent transplant.

We calculated the PAF as the fraction of deaths attributable to cancer among this cohort of transplant recipients. Specifically, $PAF = p_d \times [(\mathrm{HR} - 1)/\mathrm{HR}]$ where p_d is the proportion of all deaths preceded by cancer diagnosis, and HR is the hazard ratio that quantifies the risk of death associated with a cancer diagnosis.[28] We estimated HRs using Cox proportional hazards models with cancer as a time-dependent variable. Models were additionally adjusted for age at transplant, sex, race/ethnicity, transplanted organ, and calendar year of transplant (see Table 3 footnote). The time scale was time since transplant starting 90 days post-transplant. We calculated the cancer-attributable mortality rate as PAF × overall mortality rate (i.e., attributable deaths per 100,000 person-years). The variance of the PAF was calculated using an influence function-based approach.[64] 95% confidence intervals were computed as PAF +/- 1.96 × standard error.

We present results for all cancers combined and separately for selected cancers. We selected cancer sites that show elevated incidence among transplant recipients and were expected to have sufficient cases and subsequent deaths for the PAF calculations.[1] Breast and prostate cancers were included because these are common among the US general population. We include grouped results for cancers with infectious etiology (NHL, Hodgkin lymphoma, nasopharynx, liver, stomach, Kaposi sarcoma, anus, vulva, cervix, penis, vagina, and oropharynx including tonsil). The HR for each cancer or grouping was estimated using a separate model, in which recipients without the cancer of interest were considered unaffected. We also show PAF results for subgroups of the cohort defined by age at transplant, sex, race/ethnicity, transplanted organ, and time since transplant.

We hypothesized that differences between cancer-attributable mortality rates (calculated using PAFs) and cancer-specific mortality rates (calculated using CODs) may arise because some deaths with non-cancer COD were actually caused by cancer. To estimate this proportion, we applied our PAF methodology separately for five mortality outcomes based on different COD categories recorded in the SRTR (cancer, infection, graft failure, other, unknown). HRs were then estimated separately for each mortality outcome. For each Cox model, the events were deaths due to the specified outcome, and other individuals (including those who died of other causes) were censored at the end of follow-up. These results correspond to the proportion of deaths attributable to cancer for the specified COD category, and the sum of the five cancer-attributable mortality rates corresponds closely to the overall cancer-attributable mortality rate.

Finally, we compared CODs from the SRTR to CODs from cancer registries.[37] Analyses were restricted to deceased recipients with cancer who resided in areas covered by 12 cancer registries providing CODs. The kappa statistic was used to quantify agreement and should be close to 1.00 if both the SRTR and cancer registries accurately capture cancer-related deaths in these individuals.[65]

Computations were performed using SAS software version 9.4.

D. Results

The study included 221,962 solid organ transplant recipients (Table 2). Mean age at transplant was 45 years. The majority of recipients were male (61%) and non-Hispanic white (61%). The most commonly transplanted organs were kidney (62%), liver (18%), heart (10%), and lung (5%).

Cancer-attributable mortality

Recipients were followed for a median of 4.6 years, during which 15,012 recipients developed a first cancer and 8,123 of those with cancer died. Overall, 16.6% of deaths were preceded by a cancer diagnosis.

	N	%
Age at transplant, years		
0-17	16,707	7.5
18-34	34,997	15.8
35-49	67,595	30.5
50-64	82,092	37.0
65+	20,571	9.3
Sex		
Female	86,734	39.1
Male	135,228	60.9
Race/ethnicity		
Non-Hispanic white	134,535	60.6
Non-Hispanic black	40,114	18.1
Hispanic	33,726	15.2
Asian/Pacific islander	12,146	5.5
Other	1,441	0.7
Transplanted organ		
Kidney	136,534	61.5
Liver	39,155	17.6
Heart	22,912	10.3
Lung	10,467	4.7
Other	12,894	5.8
Year of transplant		
1987-1994	32,257	14.5
1995-1999	49,140	22.1
2000-2004	58,297	26.3
2005-2008	50,901	22.9
2009-2014	31,367	14.1

This study includes data from 17 US population-based cancer registries: California (years of cancer registry data 1988-2012), Colorado (1988-2009), Connecticut (1973- 2009), Florida (1981-2009), Georgia (1995-2010), Hawaii (1973-2007), Illinois (1986- 2013), Iowa (1973-2009), Kentucky (1995-2011), Michigan (1985-2009), New Jersey (1979-2010), New York (1976- 2010), North Carolina (1990-2010), Pennsylvania (1985-2013), Texas (1995-2010), Utah (1973-2008), and the Seattle-Puget Sound area of Washington State (1974- 2014)

Table 2: Characteristics of 221,962 solid organ transplant recipients in the United States (1987-2014)

The adjusted HR for death following a cancer diagnosis was 4.79, so the PAF for cancer-attributable mortality was $p_d \times [(\mathrm{HR} - 1)/\mathrm{HR}] = 0.166 \times [(4.79 - 1)/4.79] = 13.2\%$ (95%CI: 12.4, 13.9), corresponding to a cancer-attributable mortality rate of 516 deaths per 100,000 person-years (Table 3). Lung cancer was the largest contributor to cancer-attributable mortality (PAF=3.1%; Table 3) and had the highest cancer-attributable mortality rate (121 per 100,000 person-years). This was followed by NHL (PAF=1.9%), colorectal cancer (PAF=0.7%), and kidney cancer (PAF=0.5%).

	Cancer cases	**Deaths**	**Deaths with prior cancer, P_d (%)**	**HR***	**PAF (%)**	**95%CI**	**Cancer-attributable mortality rate**
Overall	15012	48831	16.6	4.79	13.2	(12.4, 13.9)	516
Cancer site							
Lung	2033	48831	3.4	10.42	3.1	(2.8, 3.4)	121
NHL	2240	48831	2.4	4.58	1.9	(1.6, 2.2)	75
Colorectum	877	48831	1.0	3.48	0.7	(0.6, 0.9)	29
Kidney	1184	48831	0.8	2.42	0.5	(0.3, 0.6)	18
Pancreas	273	48831	0.5	31.6	0.5	(0.4, 0.6)	19
Stomach	233	48831	0.4	14.0	0.4	(0.2, 0.5)	14
Breast	801	48831	0.6	1.79	0.3	(0.1, 0.4)	10
Bladder	345	48831	0.4	4.21	0.3	(0.2, 0.4)	12
Melanoma	581	48831	0.5	2.26	0.3	(0.2, 0.4)	11
Liver	177	48831	0.3	13.35	0.3	(0.2, 0.4)	10
Myeloma	173	48831	0.2	6.88	0.2	(0.1, 0.3)	8
Esophagus	153	48831	0.3	11.06	0.2	(0.1, 0.3)	9
Larynx	174	48831	0.2	3.27	0.1	(0.1, 0.2)	6
Thyroid	339	48831	0.1	1.40	0.0	(0.0, 0.1)	2
Prostate	1686	48831	1.0	0.99	0.0	--	-1
Infectious etiology†	3355	48831	3.7	4.81	3.0	(2.6, 3.3)	116
Non-infectious etiology	11657	48831	12.9	4.30	9.9	(9.3, 10.5)	388

Table 3: Cancer-attributable mortality among solid organ transplant recipients

Since men with prostate cancer had nominally lower risk of death than men without prostate cancer (adjusted HR=0.99), the resulting PAF was negative. Three percent of deaths were attributed to cancers of infectious etiology (cancer-attributable mortality rate 116 per 100,000 person-years), while the remaining cancers of non-infectious etiology contributed almost 10% of deaths (388 per 100,000 person-years). Results for subgroups of the cohort and individual cancers are presented in Table 4 and graphically in figures.

		Cancer cases, n	**Deaths with prior cancer, p_d (%)**	**PAF**	**95%CI**	**Cancer-attributable mortality rate**
Age						
0-17	Lung	0	0.0	0.0		0
	NHL	297	5.2	4.1	(1.0, 7.2)	71
	Colorectum	1	0.1	0.1	(-0.3, 0.4)	1
	Kidney	9	0.1	0.1	(-0.3, 0.4)	1
	Pancreas	0	0.0	0.0		0
	Stomach	0	0.0	0.0		0
	Breast	0	0.0	0.0		0
	Bladder	2	0.1	0.1	(-0.4, 0.6)	2
	Melanoma	1	0.1	0.1	(-0.3, 0.4)	1
	Liver	5	0.1	0.1	(-0.4, 0.6)	2
	All	438	7.5	6.1	(2.5, 9.8)	107
18-34	Lung	16	0.3	0.2	(-0.2, 0.7)	5
	NHL	321	3.3	2.7	(1.2, 4.1)	55
	Colorectum	31	0.3	0.3	(-0.2, 0.7)	5
	Kidney	109	0.5	0.4	(-0.2, 1.0)	8
	Pancreas	7	0.1	0.1	(-0.2, 0.5)	3
	Stomach	6	0.1	0.1	(-0.2, 0.5)	3
	Breast	64	0.5	0.4	(-0.2, 0.9)	8
	Bladder	12	0.1	0.1	(-0.2, 0.3)	2
	Melanoma	40	0.2	0.1	(-0.2, 0.4)	2
	Liver	8	0.2	0.2	(-0.2, 0.5)	3
	All	1017	8.9	7.3	(4.9, 9.7)	151

Table 4: Cancer-attributable mortality rates, according to individual and transplant characteristics and cancer site

Table 4 continued

		Cancer cases, n	Deaths with prior cancer, p_d (%)	PAF	95%CI	Cancer-attributable mortality rate
35-49	Lung	347	2.1	2.0	(1.4, 2.5)	62
	NHL	562	2.3	1.8	(1.2, 2.4)	57
	Colorectum	206	0.8	0.6	(0.3, 0.9)	18
	Kidney	430	0.8	0.4	(0.1, 0.8)	14
	Pancreas	54	0.4	0.3	(0.1, 0.6)	11
	Stomach	55	0.3	0.3	(0.1, 0.5)	10
	Breast	262	0.6	0.4	(0.1, 0.6)	11
	Bladder	63	0.2	0.2	(0.0, 0.4)	6
	Melanoma	137	0.4	0.3	(0.0, 0.5)	9
	Liver	49	0.3	0.3	(0.1, 0.5)	8
	All	3637	13.1	10.5	(9.1, 11.9)	329
50-64	Lung	1319	4.8	4.3	(3.8, 4.8)	228
	NHL	832	2.2	1.7	(1.4, 2.0)	90
	Colorectum	468	1.2	0.8	(0.6, 1.1)	44
	Kidney	514	0.9	0.5	(0.3, 0.7)	26
	Pancreas	146	0.6	0.6	(0.4, 0.7)	29
	Stomach	139	0.5	0.4	(0.3, 0.6)	23
	Breast	388	0.6	0.2	(0.0, 0.4)	11
	Bladder	201	0.5	0.4	(0.2, 0.6)	21
	Melanoma	300	0.6	0.3	(0.2, 0.5)	17
	Liver	95	0.3	0.3	(0.2, 0.4)	16
	All	7736	19.7	15.5	(14.5, 16.5)	818
65+	Lung	351	4.2	3.7	(2.9, 4.4)	292
	NHL	228	2.2	1.7	(1.1, 2.2)	134
	Colorectum	171	1.6	1.1	(0.6, 1.5)	87
	Kidney	122	0.9	0.6	(0.2, 0.9)	44
	Pancreas	66	0.9	0.8	(0.5, 1.2)	68
	Stomach	33	0.4	0.4	(0.1, 0.6)	28
	Breast	87	0.5	0.2	(-0.1, 0.4)	14
	Bladder	67	0.6	0.4	(0.2, 0.7)	35
	Melanoma	103	0.7	0.3	(0.0, 0.6)	25
	Liver	20	0.2	0.2	(0.0, 0.4)	17
	All	2185	20.3	15.5	(13.8, 17.1)	1229

Table 4 continued

		Cancer cases, n	Deaths with prior cancer, p_d (%)	PAF	95%CI	Cancer-attributable mortality rate
Organ Transplanted						
Kidney	Lung	891	3.2	3.0	(2.4, 3.5)	93
	NHL	1039	2.1	1.7	(1.2, 2.1)	53
	Colorectum	440	1.1	0.8	(0.5, 1.1)	25
	Kidney	909	1.1	0.6	(0.3, 0.9)	19
	Pancreas	157	0.6	0.6	(0.3, 0.8)	18
	Stomach	123	0.4	0.4	(0.2, 0.6)	12
	Bladder	196	0.5	0.4	(0.2, 0.6)	12
	Melanoma	335	0.6	0.3	(0.1, 0.5)	10
	Liver	82	0.3	0.3	(0.1, 0.4)	8
	All	8029	16.6	13.1	(11.9, 14.3)	410
Liver	Lung	397	3.2	2.9	(2.2, 3.5)	118
	NHL	488	2.4	1.8	(1.2, 2.4)	74
	Colorectum	221	1.2	0.9	(0.5, 1.3)	36
	Kidney	77	0.3	0.2	(0.0, 0.4)	8
	Pancreas	63	0.6	0.6	(0.3, 0.9)	24
	Stomach	48	0.4	0.4	(0.1, 0.6)	15
	Bladder	45	0.3	0.2	(0.0, 0.4)	9
	Melanoma	84	0.3	0.2	(0.0, 0.4)	7
	Liver	66	0.5	0.5	(0.2, 0.7)	19
	All	2781	15.8	13.0	(11.6, 14.5)	531
Heart	Lung	496	5.1	4.6	(3.8, 5.4)	251
	NHL	391	3.2	2.4	(1.8, 3.0)	132
	Colorectum	114	1.0	0.7	(0.4, 1.0)	39
	Kidney	126	0.8	0.4	(0.1, 0.7)	23
	Pancreas	33	0.3	0.3	(0.1, 0.5)	18
	Stomach	40	0.4	0.4	(0.2, 0.6)	22
	Bladder	64	0.5	0.3	(0.1, 0.5)	18
	Melanoma	118	0.8	0.4	(0.2, 0.7)	23
	Liver	20	0.2	0.2	(0.0, 0.4)	11
	All	2628	20.9	16.4	(14.8, 17.9)	891
Lung	Lung	207	3.5	3.0	(2.3, 3.6)	349
	NHL	162	2.4	1.7	(1.1, 2.2)	198
	Colorectum	67	0.9	0.5	(0.2, 0.8)	59
	Kidney	27	0.4	0.3	(0.1, 0.5)	36
	Pancreas	8	0.2	0.1	(0.0, 0.3)	18
	Stomach	16	0.2	0.2	(0.0, 0.4)	24
	Bladder	22	0.3	0.2	(0.0, 0.4)	27
	Melanoma	22	0.3	0.1	(-0.1, 0.3)	14
	Liver	5	0.1	0.0	(-0.1, 0.1)	5

Table 4 continued

		Cancer cases, n	Deaths with prior cancer, p_d (%)	PAF	95%CI	Cancer-attributable mortality rate
	All	942	13.9	10.5	(9.2, 11.8)	1241
Other	Lung	42	1.1	1.0	(0.2, 1.9)	37
	NHL	160	3.6	3.1	(1.6, 4.6)	108
	Colorectum	35	0.6	0.4	(-0.2, 1.0)	15
	Kidney	45	0.7	0.5	(-0.2, 1.1)	16
	Pancreas	12	0.4	0.4	(-0.1, 0.9)	14
	Stomach	6	0.2	0.1	(-0.2, 0.5)	5
	Bladder	18	0.3	0.3	(-0.2, 0.7)	9
	Melanoma	22	0.3	0.3	(-0.2, 0.7)	9
	Liver	4	0.1	0.1	(-0.2, 0.4)	4
	All	633	11.8	9.8	(7.0, 12.6)	342
Time since transplant, years						
0-1.99	Lung	464	1.9	1.8	(1.0, 2.6)	68
	NHL	869	2.3	2.1	(1.2, 3.0)	80
	Colorectum	200	0.3	0.3	(-0.1, 0.6)	11
	Kidney	325	0.3	0.3	(-0.1, 0.6)	10
	Pancreas	74	0.4	0.3	(0.0, 0.7)	13
	Stomach	49	0.2	0.1	(-0.1, 0.4)	6
	Breast	193	0.1	0.1	(-0.1, 0.3)	3
	Bladder	70	0.2	0.2	(-0.1, 0.4)	6
	Melanoma	187	0.2	0.2	(-0.1, 0.5)	7
	Liver	39	0.1	0.1	(-0.1, 0.4)	5
	All	4159	8.9	8.1	(6.4, 9.9)	313
2-4.99	Lung	825	4.0	3.7	(2.6, 4.9)	126
	NHL	939	1.8	1.5	(0.7, 2.2)	50
	Colorectum	401	1.0	0.8	(0.2, 1.3)	26
	Kidney	548	0.8	0.5	(0.0, 1.0)	18
	Pancreas	101	0.6	0.6	(0.1, 1.0)	19
	Stomach	94	0.5	0.4	(0.0, 0.8)	14
	Breast	401	0.5	0.2	(-0.1, 0.6)	8
	Bladder	132	0.3	0.2	(-0.1, 0.5)	7
	Melanoma	317	0.5	0.3	(-0.1, 0.7)	9
	Liver	75	0.3	0.3	(0.0, 0.7)	11
	All	6864	16.1	13.5	(11.3, 15.7)	456
5-9.99	Lung	887	4.3	3.8	(2.9, 4.7)	159
	NHL	1116	2.9	2.2	(1.5, 2.9)	92
	Colorectum	489	1.5	1.0	(0.5, 1.5)	42
	Kidney	649	0.9	0.5	(0.1, 0.9)	20
	Pancreas	96	0.6	0.6	(0.2, 0.9)	24
	Stomach	105	0.5	0.5	(0.2, 0.8)	19

Table 4 continued

		Cancer cases, n	Deaths with prior cancer, p_d (%)	PAF	95%CI	Cancer-attributable mortality rate
	Breast	517	0.8	0.3	(0.0, 0.7)	15
	Bladder	212	0.7	0.5	(0.2, 0.9)	22
	Melanoma	344	0.7	0.4	(0.0, 0.7)	15
	Liver	69	0.3	0.3	(0.1, 0.6)	13
	All	8018	21.3	16.0	(14.1, 18.0)	673
10+	Lung	396	4.0	3.3	(2.2, 4.3)	168
	NHL	645	3.3	1.9	(1.0, 2.9)	99
	Colorectum	290	1.9	1.1	(0.4, 1.8)	56
	Kidney	403	1.6	0.7	(0.1, 1.4)	36
	Pancreas	37	0.5	0.5	(0.1, 0.9)	24
	Stomach	39	0.5	0.4	(0.0, 0.8)	22
	Breast	308	1.2	0.4	(-0.2, 0.9)	19
	Bladder	87	0.6	0.4	(0.0, 0.8)	20
	Melanoma	171	0.9	0.3	(-0.2, 0.7)	14
	Liver	36	0.3	0.3	(0.0, 0.6)	15
	All	4411	25.5	15.7	(13.0, 18.4)	810
Race/ethnicity						
Non-Hispanic white	Lung	1602	3.9	3.5	(3.1, 3.9)	149
	NHL	1656	2.7	2.1	(1.7, 2.4)	87
	Colorectum	629	1.1	0.8	(0.6, 1.0)	32
	Kidney	612	0.7	0.4	(0.2, 0.6)	17
	Pancreas	166	0.4	0.4	(0.3, 0.6)	18
	Stomach	146	0.4	0.3	(0.2, 0.5)	14
	Breast	529	0.5	0.2	(0.1, 0.3)	8
	Bladder	282	0.5	0.4	(0.2, 0.5)	15
	Melanoma	556	0.7	0.4	(0.2, 0.5)	16
	Liver	106	0.3	0.2	(0.1, 0.3)	10
	All	10755	18.0	14.2	(13.4, 15.1)	600
Non-Hispanic black	Lung	275	2.7	2.5	(1.7, 3.3)	104
	NHL	204	1.2	0.9	(0.4, 1.4)	38
	Colorectum	138	1.0	0.7	(0.3, 1.2)	31
	Kidney	366	1.3	0.7	(0.2, 1.2)	30
	Pancreas	54	0.6	0.5	(0.2, 0.9)	22
	Stomach	45	0.4	0.4	(0.1, 0.7)	15
	Breast	147	0.7	0.4	(0.1, 0.8)	18
	Bladder	25	0.2	0.1	(-0.1, 0.3)	5
	Melanoma	7	0.0	0.0	(0.0, 0.0)	0
	Liver	32	0.3	0.3	(0.0, 0.5)	12
	All	2241	13.4	10.2	(8.5, 11.8)	423

Table 4 continued

		Cancer cases, n	Deaths with prior cancer, p_d (%)	PAF	95%CI	Cancer-attributable mortality rate
Hispanic	Lung	91	1.5	1.4	(0.5, 2.2)	39
	NHL	258	2.6	2.2	(1.1, 3.3)	64
	Colorectum	68	0.7	0.5	(0.0, 1.0)	14
	Kidney	157	0.8	0.5	(-0.1, 1.1)	15
	Pancreas	38	0.7	0.7	(0.1, 1.2)	19
	Stomach	18	0.3	0.3	(-0.1, 0.7)	9
	Breast	82	0.5	0.4	(-0.1, 0.8)	10
	Bladder	30	0.4	0.4	(-0.1, 0.8)	10
	Melanoma	13	0.2	0.2	(-0.1, 0.5)	5
	Liver	30	0.4	0.4	(-0.1, 0.8)	10
	All	1411	13.0	10.7	(8.3, 13.0)	303
Asian/Pacific islander	Lung	51	2.5	2.3	(0.4, 4.3)	59
	NHL	105	3.2	2.8	(0.5, 5.0)	70
	Colorectum	41	1.3	1.0	(-0.4, 2.4)	25
	Kidney	37	0.7	0.4	(-0.6, 1.4)	10
	Pancreas	13	0.7	0.7	(-0.4, 1.8)	18
	Stomach	24	1.0	1.0	(-0.3, 2.2)	24
	Breast	40	0.5	0.4	(-0.5, 1.2)	9
	Bladder	8	0.1	0.1	(-0.3, 0.5)	2
	Melanoma	4	0.0	0.0		0
	Liver	9	0.3	0.3	(-0.4, 1.0)	7
	All	531	16.0	13.8	(9.0, 18.6)	350
Other	Lung	14	4.4	4.1	(-0.8, 9.0)	168
	NHL	17	3.0	2.2	(-2.0, 6.5)	92
	Colorectum	1	0.3	0.3	(-1.1, 1.8)	14
	Kidney	12	0.7	0.2	(-1.8, 2.3)	10
	Pancreas	2	0.7	0.7	(-1.4, 2.7)	27
	Stomach	0	0.0	0.0		0
	Breast	3	0.3	1.3	(-322, 325)	51
	Bladder	0	0.0	0.0		0
	Melanoma	1	0.0	0.0		0
	Liver	0	0.0	0.0		0
	All	75	14.8	12.4	(3.4, 21.4)	506

Table 4 continued

		Cancer cases, n	Deaths with prior cancer, p_d (%)	PAF	95%CI	Cancer-attributable mortality rate
Sex						
Female	Lung	645	2.8	2.5	(2.0, 3.0)	91
	NHL	764	2.1	1.6	(1.2, 2.1)	59
	Colorectum	359	1.1	0.8	(0.5, 1.1)	28
	Kidney	333	0.6	0.4	(0.1, 0.6)	13
	Pancreas	97	0.5	0.5	(0.3, 0.7)	17
	Stomach	67	0.3	0.2	(0.1, 0.4)	8
	Breast	769	1.5	0.7	(0.3, 1.0)	24
	Bladder	83	0.3	0.3	(0.1, 0.4)	9
	Melanoma	154	0.3	0.1	(0.0, 0.3)	4
	Liver	33	0.1	0.1	(0.0, 0.2)	4
	All	4972	14.1	11.1	(10.0, 12.2)	404
Male	Lung	1388	3.8	3.4	(3.0, 3.9)	141
	NHL	1476	2.6	2.1	(1.7, 2.4)	86
	Colorectum	518	1.0	0.7	(0.5, 0.9)	30
	Kidney	851	0.9	0.5	(0.3, 0.7)	22
	Pancreas	176	0.5	0.5	(0.3, 0.7)	20
	Stomach	166	0.5	0.4	(0.3, 0.6)	17
	Breast	32	0.0	0.0	(0.0, 0.1)	1
	Bladder	262	0.5	0.3	(0.2, 0.5)	14
	Melanoma	427	0.6	0.4	(0.2, 0.5)	16
	Liver	144	0.4	0.3	(0.2, 0.5)	14
	All	10041	18.1	14.4	(13.4, 15.3)	590

Overall and cancer-attributable mortality rates are per 100,000 person-years. Cancers are ordered by overall PAF.

PAFs increased with age at transplant, from 6.1% of deaths among children to 15.5% of deaths among recipients aged 50+ years (Table 3, Figure 3). Cancer-attributable mortality rates also increased steeply with age, reaching 1229 per 100,000 person-years among recipients 65+ years old at transplant. In the youngest age group, NHL was the largest single contributor to cancer-attributable mortality (PAF=4.1%, Figure 3). In the oldest age groups, however, lung cancer was the largest contributor (PAFs=3.7-4.3%

among 50+ year-olds) along with NHL (1.7%), kidney cancer (0.5-0.6%), and colorectal cancer (0.8-1.1%).

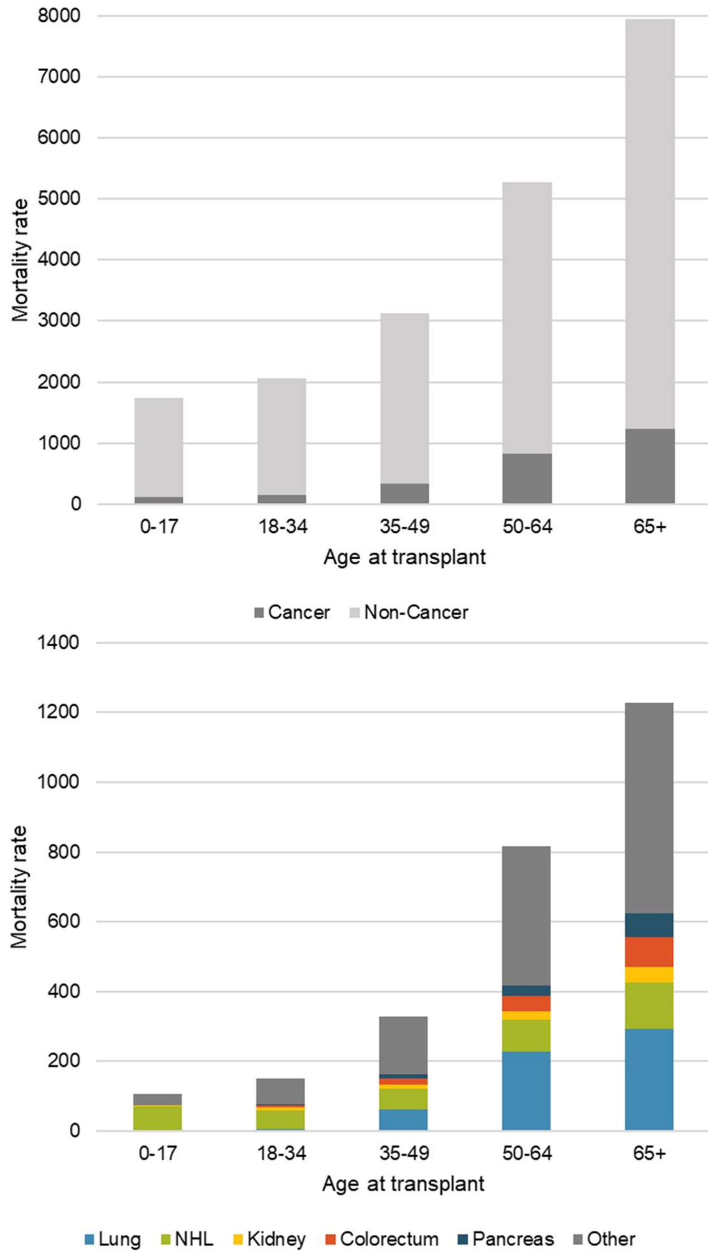


Figure 3: Cancer-attributable mortality rates, according to age at transplant

The top panel displays cancer and non-cancer mortality rates per 100,000 person-years according to age at transplant. The bottom panel displays cancer-attributable mortality rates per 100,000 person-years according to the age at transplant for each cancer site.

The PAF was higher among males than females (14.4% vs. 11.1%, Table 3). Among males, the largest contributor was lung cancer (PAF=3.4%) followed by NHL (PAF=2.1%). Although these cancers were also the largest contributors in females, their contribution was lower (PAF=2.5% and 1.6%, respectively) (Table 4, Figure 4).

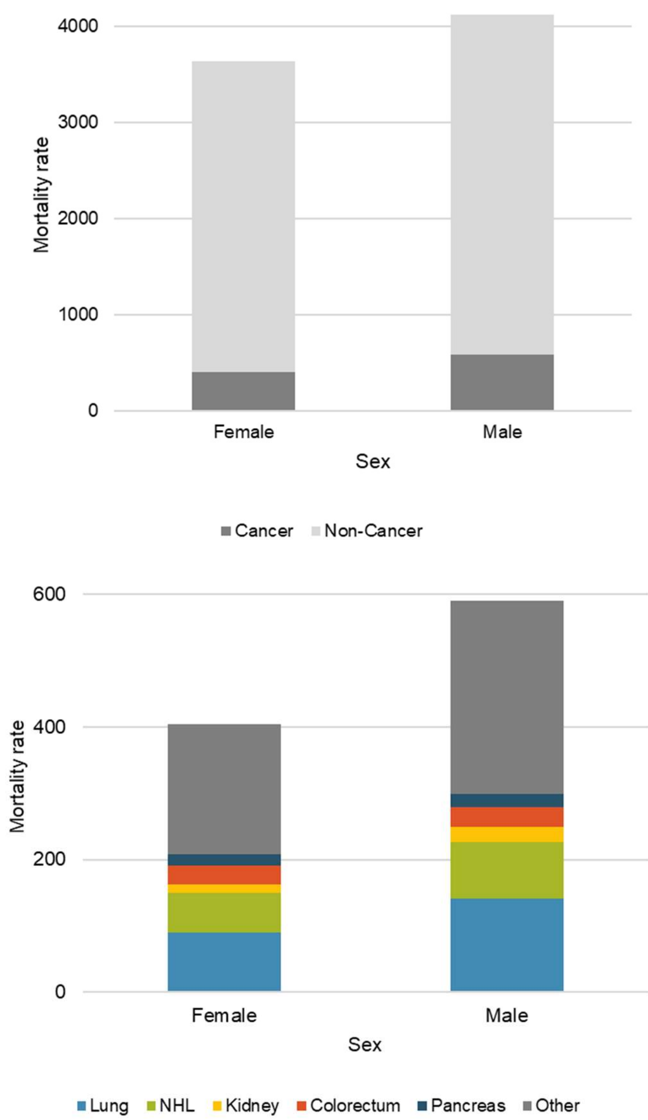


Figure 4: Cancer-attributable mortality rates, according to sex

The top panel displays cancer and non-cancer mortality rates per 100,000 person-years according to sex. The right panel displays cancer-attributable mortality rates per 100,000 person-years according to sex for each cancer site.

PAFs varied by race/ethnicity with non-Hispanic whites having the highest PAF (14.2%), followed by Asian/Pacific Islanders (13.8%). Non-Hispanic blacks and Hispanics had lower PAFs (10.2-10.7%). Although there was some variability, lung cancer and NHL made the greatest contributions across the racial/ethnic groups (PAFs 1.4-4.1% and 0.9-2.8% respectively; Table 4, Figure 5).

PAFs varied by transplanted organ, from 16.4% among heart recipients to 9.8% among recipients of other (miscellaneous) organs (Figure 6 and Table 3). Lung recipients had the highest overall mortality and thus the highest cancer-attributable mortality rate (1241 per 100,000 person-years). Heart recipients had the next highest cancer-attributable mortality rate (891 per 100,000 person-years), followed by liver recipients (531 per 100,000 person-years). NHL and lung cancer were major contributors across all organs (PAF=1.7-3.1% for NHL, 1.0-4.6% for lung cancer). The proportion of deaths attributed to kidney cancer among kidney recipients was low (PAF=0.6%). The PAF for liver cancer was small among liver recipients (0.5%), however, this was almost double compared to recipients of other organs (0.1%-0.3%, Figure 6 and Table 4).

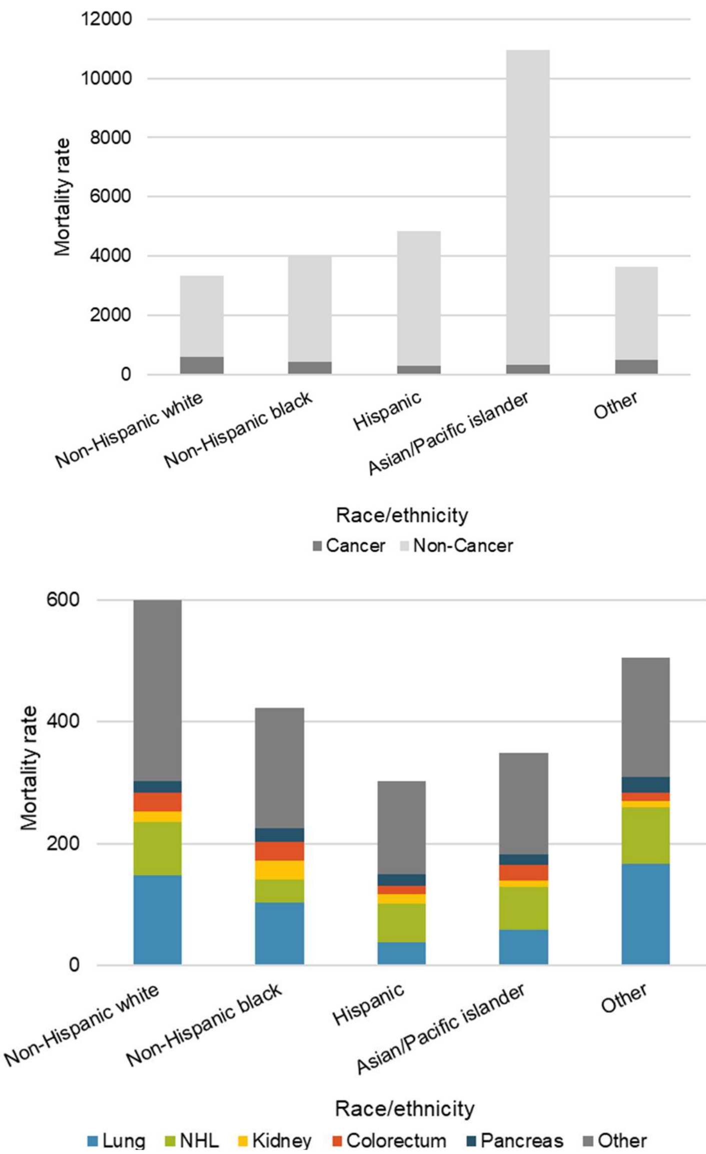


Figure 5: Cancer-attributable mortality rates, according to race/ethnicity

The top panel displays cancer and non-cancer mortality rates per 100,000 person-years according to race/ethnicity. The bottom panel displays cancer-attributable mortality rates per 100,000 person-years according to race/ethnicity for each cancer site

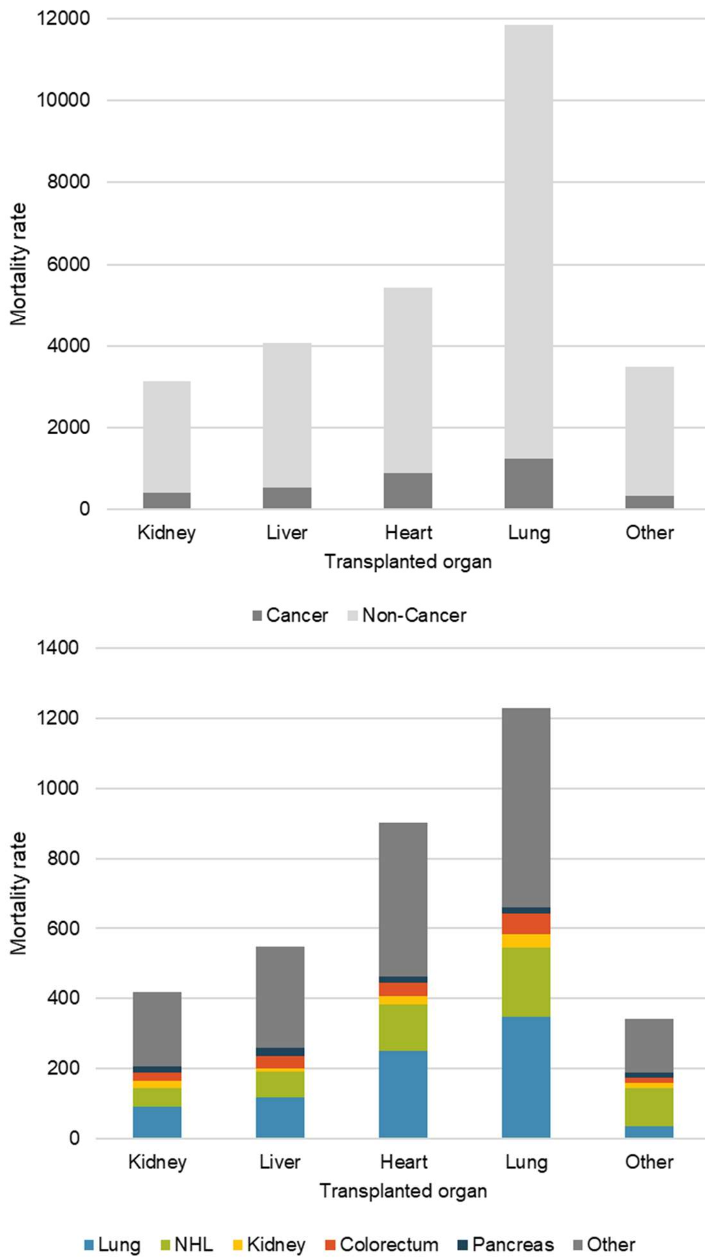


Figure 6: Cancer-attributable mortality rates, according to transplanted organ

The top panel displays cancer and non-cancer mortality rates per 100,000 person-years according to transplanted organ. The bottom panel displays cancer-attributable mortality rates per 100,000 person-years according to the transplanted organ for each cancer site.

Overall mortality was high 0-2 years post-transplant, decreased 2-5 years post-transplant, and then increased subsequently (Figure 7 and Table 3). PAFs increased over time, from 8.1% (0-2 years post-transplant) to 16.0% and 15.7% (5-10 years and 10+ years post-transplant, respectively). Thus, the cancer-attributable mortality rate increased over time, reaching 810 per 100,000 person-years during 10+ years post-transplant. For lung cancer, PAFs increased from 1.8% within 2 years of transplant to approximately 4% for 2+ years post-transplant, and cancer-attributable mortality for lung cancer increased over time from 68 to 168 per 100,000 person-years (Figure 7). The PAF for NHL was 2.1% during 0-2 years post-transplant, decreased slightly to 1.5% during 2-5 years post-transplant, then increased again to approximately 2% for 5+ years post-transplant (Figure 7 and Table 4).

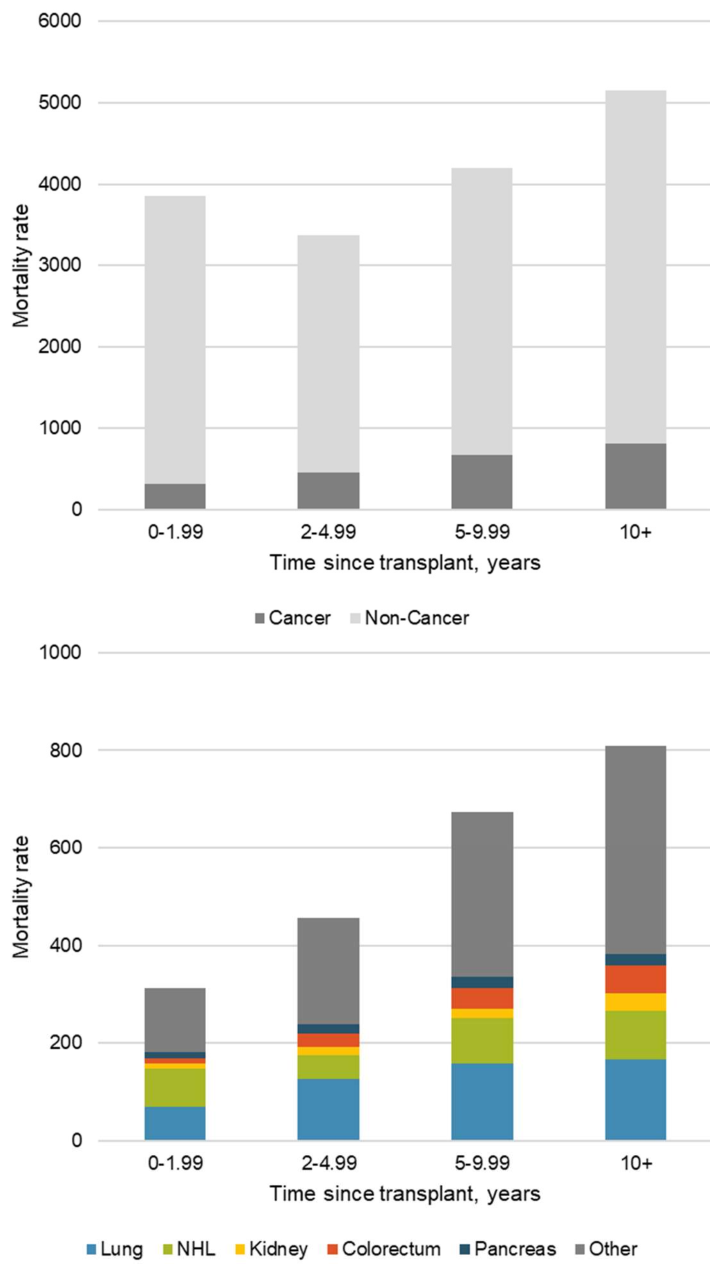


Figure 7: Cancer-attributable mortality rates, according to time since transplant

The top panel displays cancer and non-cancer mortality rates per 100,000 person-years according to years since transplant. The bottom panel displays cancer-attributable mortality rates per 100,000 person-years according to years since transplant for each cancer site.

Comparison of cancer-attributable and cancer-specific mortality rates

Based on SRTR CODs, we calculated a cancer-specific mortality rate of 368 per 100,000 person-years, substantially lower than the above cancer-attributable mortality rate (516 per 100,000 person-years). PAFs for each of five mortality outcomes were assessed separately. With COD specified as cancer, a high proportion of deaths (PAF=74%) were attributed to cancer, as expected (Table 5). Notably, additional deaths were also attributed to cancer among deaths with other CODs. Specifically, we attributed to cancer 4.2% of deaths with infection recorded as the COD, 4.2% of deaths with graft failure as the COD, 5.6% of deaths with other recorded CODs, and 9.4% of unknown CODs. The sum of the cancer-attributable mortality rates across the five COD categories was 491 per 100,000 person-years, close to the overall cancer-attributable mortality rate (516 per 100,000 person-years).

Mortality outcome category defined by COD in SRTR	Deaths	Mortality rate	Deaths with prior cancer, P_d (%)	HR*	PAF (%)	Cancer-attributable mortality rate
Cancer	4584	368	74.4	138	73.8	272
Infection	6683	537	7.4	2.3	4.2	23
Graft failure	6001	482	7.5	2.3	4.2	20
Other causes	20154	1619	9.7	2.4	5.6	90
Missing	11409	916	13.8	3.2	9.4	86
Overall	48831	--	--	--	--	491

HR=Hazard ratio; PAF=Population-attributable fraction

*Hazard ratios are computed using time since transplant as the time scale and are adjusted for sex, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian/Pacific Islander, other), organ transplanted (kidney, liver, heart, lung, other), calendar year of transplant (1987-1994, 1995-1999, 2000-2004, 2005-2008, 2009-2014) and age at transplant. Since the effect of age on mortality was non-linear, age was modeled as a piecewise linear function with knots at age 5 and 19 years.

Overall and cancer-attributable mortality rates are per 100,000 person-years

Table 5: PAF and cancer-attributable mortality for five mortality outcomes

Finally, we conducted an analysis restricted to 12 cancer registries that provided CODs for cancer cases (N=165,839). In this subgroup, approximately 25% of deceased recipients with cancer had missing or unknown CODs in the SRTR or cancer registry. There were 4943 recipients with a cancer diagnosis who subsequently died and had CODs specified in both the SRTR and cancer registry. Of these, 2773 (56%) had cancer listed as the COD in the SRTR and 2807 (57%) had cancer listed as the COD in the cancer registry. However, only 2062 (42%) had cancer listed as the COD in both the SRTR and cancer registry (kappa=0.40).

E. Discussion

Among US solid organ transplant recipients, 13.2% of deaths were attributed to cancer, of which almost half were attributable to NHL and lung, kidney, and colorectal cancers. The PAF depends on cancer incidence and its impact on mortality. The two cancer sites with the highest PAFs (lung cancer and NHL) are both common and associated with high mortality in transplant recipients. Moreover, transplant recipients have worse survival following cancer diagnosis than cancer patients in the general population,[62] contributing further to cancer-attributable mortality.

Immunosuppression plays a major role in malignancy after transplant.[18] Indeed, we found that lung and heart recipients, who generally receive intensive immunosuppression, had the highest cancer-attributable mortality. Lung and heart recipients had high PAFs for NHL, a cancer strongly associated with EBV, and for lung cancer, which is linked to smoking and, among lung recipients, end-stage lung disease.[17] Among the overall cohort, the cancer-attributable mortality rate also increased with time since transplant, and lung cancer and NHL contributed prominently greater than five years post-transplant. This pattern partly reflects rising incidence of these cancers with longer duration since transplant and prolonged exposure to immunosuppressant medications.[1,66,67]

Among children, NHL was the single largest contributor to cancer-attributable mortality. NHL, when it occurs in transplant recipients, is a type of post-transplant lymphoproliferative disorder (PTLD), which among young recipients results from primary EBV infection.[68-70] Other cancers such as lung and colorectal cancers become more common with age and so contribute more to mortality in older recipients. In addition, the PAF was higher among males compared to females for all cancer sites combined and for individual cancers, although the relative ranking of cancer sites was similar. While there were some differences in PAFs among the racial/ethnic groups, lung cancer and NHL comprised the largest contributors for each group.

PAF calculations rely on a causal relationship between cancer and excess mortality, and a key assumption is that the association between cancer and death is estimated without confounding. Our Cox models were adjusted for recipient demographic characteristics. However, if there was substantial residual confounding, then PAFs would not accurately reflect mortality that should be attributed to cancer. For example, tobacco use could be a strong confounder of the association between lung cancer and mortality. Since the SRTR does not have complete information on tobacco use, we calculated a bias factor to adjust the HR for smoking (see Statistical details).[71] This method relies on estimates of the relative risk of death for smokers vs. non-smokers and smoking prevalence among recipients. Our calculations yield a bias factor of 1.67 which, when applied to lung cancer, resulted in an HR further adjusted for smoking of 6.28. Although this is lower than the HR of 10.42 used in our primary analyses, it results in a similar PAF (2.5% vs. 3.1%). Therefore, we conclude that the PAF for lung cancer is not greatly affected by unmeasured confounding by tobacco use.

Confounding likely explains the slightly negative PAF for prostate cancer since it is implausible that prostate cancer lowers the risk of death. Most prostate cancer cases are diagnosed through screening, and many detected cancers are indolent.[72,73] The lower mortality observed in our study among recipients with prostate cancer thus probably reflects the relative health of men offered screening.

The cancer-specific mortality rate based on CODs from the SRTR was substantially lower than our cancer-attributable mortality rate estimate (368 vs. 516 per 100,000 person-years). We believe this reflects inaccurate determination of CODs, because recipients may appear to die from non-cancer CODs, such as infection or graft failure, that are actually downstream effects of cancer or its treatment. Indeed, Table 3 shows 4% of deaths with CODs recorded as infection and 4% as graft failure were attributable to recipients' cancer diagnoses. In addition, 5% of deaths with other causes and 9% of deaths of unknown cause were attributable to cancer, and numerically these contributed substantially to the overall cancer-attributable mortality rate. Furthermore, our comparison of CODs from the SRTR and cancer registries yielded only modest agreement (kappa=0.40), indicating that assigning CODs is difficult and that likely neither source is entirely accurate in distinguishing cancer from non-cancer deaths.

Strengths of our study include a large sample representative of the US transplant population, including recipients of all organ types. We obtained incident cancer information from population-based cancer registries that ascertain cancers in their catchment areas. We assessed potential confounding of PAFs and compared our results with cancer-specific mortality rates derived from CODs. A limitation is that cancer registries do not collect information on non-melanoma skin cancers. While these cause substantial morbidity among transplant recipients,[56,57] they are not often fatal and therefore would not greatly contribute to mortality. Also, some cancers may have been missed by the cancer registries, if the linkage did not match the cancer to the recipient, or if recipients migrated from the catchment area. However, outmigration is uncommon.[1]

There are opportunities to reduce cancer mortality in transplant recipients through prevention and screening. Most cancer screening guidelines for recipients follow recommendations for the general population, but our results indicate that more tailored guidelines would be useful.[48] Given the high PAF for NHL (4% among children and about 2% among adults), this cancer should be a priority. Since high levels of circulating EBV load can be used as a marker of PTLD, there may be opportunities to use this biomarker to screen for NHL, especially in pediatric recipients.[74] Treating underlying liver disease (e.g., with direct-acting antiviral medications for hepatitis C virus infection) and screening for liver cancers in high-risk recipients should also be considered,[48,75] including among liver recipients, among whom liver cancer contributes the most to mortality. Finally, cancer prevention strategies aimed at modifying lifestyle cancer risk factors, especially smoking cessation but also sun protection,[76] should be prioritized to reduce mortality from lung cancer and melanoma.

In conclusion, cancer causes substantial mortality among transplant recipients. About 13% of deaths are attributable to cancer, with about 5% of deaths from NHL and lung cancer combined. Cancer-attributable mortality increases with age and time since transplant, indicating that deaths from cancer will increase in the future as survival following transplantation improves. Further research aimed at cancer prevention, screening, and treatment will be required.

Statistical details

I calculated the smoking-adjusted HR using the method of Steenland and Greenland.[77] Since smoking is hypothesized to be the strongest potential confounder of the association between lung cancer and mortality, I applied this methodology to smoking as a potential confounder of the lung cancer estimate.

Let "exp" denote transplant recipients who developed an incident cancer and "nonexp" denote those who did not develop lung cancer. Let I_{exp}, $I_{exp,smk}$ and $I_{exp,nonsmk}$ be the mortality rate among those with lung cancer in the transplant population overall, and among smokers and non-smokers, respectively and let $p_{exp,smk}$ and $p_{exp,nonsmk}$ be the proportions of current smokers and non-smokers among transplant recipients with lung cancer. Let I_{nonexp}, $I_{nonexp,smk}$ and $I_{nonexp,nonsmk}$ be the mortality rate among those without lung cancer in the transplant population overall, and among smokers and non-smokers, respectively and let $p_{nexp,smk}$ and $p_{nexp,nonsmk}$ be the proportions of current smokers and non-smokers among transplant recipients without lung cancer. Then

$$HR_{unadjusted} = \frac{I_{exp}}{I_{nonex}} = \frac{\text{Mortality rate among recipients with lung cancer}}{\text{Mortality rate among recipients without lung cancer}}$$

$$HR_{adjusted} = \frac{I_{exp,nonsmk}}{I_{nonexp,nonsmk}} = \frac{\text{Mortality rate among non-smoking recipients with lung cancer}}{\text{Mortality rate among non-smoking recipients without lung cancer}}$$

$$I_{exp} = I_{exp,smk} * p_{exp,smk} + I_{exp,nonsmk} * p_{exp,nonsmk}$$
$$I_{nonexp} = I_{nonexp,smk} * p_{nonexp,smk} + I_{nonexp,nonsmk} * p_{nonexp,nonsmk}$$

Let RR_{smk} be the RR for death among smokers vs. non-smokers in the transplant population. Assume the same RR between smoking and death for transplant recipients with lung cancer and those without.

$$I_{exp} = RR_{smk} * I_{exp,nonsmk} * p_{exp,smk} + I_{exp,nonsmk} * p_{exp,nonsmk} = I_{exp,nonsmk}\left[RR_{smk} * p_{exp,smk} + p_{exp,nonsmk}\right]$$

$$I_{nonexp} = RR_{smk} * I_{nonexp,nonsmk} * p_{nonexp,smk} + I_{nonexp,nonsmk} * p_{nonexp,nonsmk} = I_{nonexp,nonsmk}\left[RR_{smk} * p_{nonexp,smk} + p_{nonexp,nonsmk}\right]$$

$$HR_{unadjusted} = \frac{I_{exp}}{I_{nonexp}} = \frac{I_{exp,nonsmk}\left[RR_{smk} * p_{exp,smk} + p_{exp,nonsmk}\right]}{I_{nonexp,nonsmk}\left[RR_{smk} * p_{non\quad,smk} + p_{nonexp,nonsmk}\right]}$$
$$HR_{unadjusted} = HR_{adjusted} * \left[\frac{RR_{smk} * p_{exp,smk} + p_{exp,nonsmk}}{RR_{smk} * p_{nonexp,smk} + p_{nonexp,nonsmk}}\right]$$

$$HR_{adjusted} = HR_{unadjusted} * \left[\frac{RR_{smk} * p_{exp,smk} + p_{exp,nonsmk}}{RR_{smk} * p_{nonexp,smk} + p_{nonexp,nonsmk}}\right]^{-1}$$

The term inside the brackets in the last two lines above is termed the “bias factor.” The parameters used to estimate the bias factor are shown in Table 6.

Parameter	Estimate
RR_{smk}: Relative risk of death for current vs. never smokers in the transplant population[78-80]	2
$p_{nonexp,smk}$: prevalence of smoking among transplant recipients without incident lung cancer[78-80]	Range 10-30%; Assume 20%
$p_{exp,smk}$: prevalence of smoking among transplant recipients with incident lung cancer	Assume 100% to maximize the bias factor correction

Table 6: Parameter estimates used in bias factor analysis

5. Cancer-attributable mortality using cause of death and comparison of results (Aim 1.2 and 1.3)

The cancer-attributable mortality rate using the PAF method was higher than when calculated using COD from the SRTR (516 vs. 368 per 100,000 person years; Chapter 4). This difference was investigated in several ways using data from cancer registries that provided COD information. These were the following 12 registries: California, Colorado, Connecticut, Georgia, Iowa, Illinois, Kentucky, New Jersey, Pennsylvania, Seattle, Texas, Michigan. COD was not available from five cancer registries: North Carolina, Florida, Hawaii, New York and Utah. Consequently, these states were excluded from the estimation of cancer mortality using COD from the cancer registries leaving 165,839 recipients for analysis (75% of the full cohort). Note that all results shown in this Chapter are using data from the 12 cancer registries that provided COD data and, therefore, are not exactly the same as those shown in Chapter 4.

First I assessed if the estimates produced using the PAF may be biased (Section A) and the second I examined the quality of the COD coding (Section B).

A. Potential impact of unmeasured confounding on the hazard ratio used to calculate the population attributable fraction

Estimates of the cancer-attributable mortality rate using the PAF rely on an estimate of the HR that is unconfounded. Further analysis was conducted in addition to the bias factor analysis performed for the impact of smoking on lung cancer in Chapter 4.

1) Estimate what the hazard ratio would need to be so that the cancer-attributable mortality rate from the PAF method equals that from the COD method

The hazard ratio necessary to produce the same cancer mortality rate that was estimated using COD can be found with the algebra below. It provides insight into whether there may be a confounder or set of confounders that could plausibly reduce the HR sufficiently to make the rates more similar.

Let D_{total} be the number of deaths in the cohort and D_{cancer} be the number of deaths among those with cancer in the cohort. Then the percentage of deaths with a prior cancer diagnosis is:

$$p_d = \frac{D_{cancer}}{D_{total}} = \frac{6144}{36749} = 0.167$$

The overall mortality rate is:

$$Rate_{Overall} = \frac{D_{Total}}{Person\ Years} = \frac{36749}{938084} = 0.0392$$

The PAF associated with cancer can be calculated by the following:

$$PAF = p_d \times \left(\frac{HR - 1}{HR}\right)$$

where p_d is the percentage of deaths with a prior cancer diagnosis among all deaths in the cohort and HR is the estimated hazard ratio for cancer.

The cancer-attributable mortality rate can be estimated by:

$$Rate_{Cancer;\,PAF} = PAF \times Rate_{Overall}$$

$$= p_d \times \left(\frac{HR-1}{HR}\right) \times Rate_{Overall}$$

Let $D_{cancer;\,COD}$ be the number of cancer deaths in the cohort as classified using the COD algorithm.

Then the cancer mortality rate can be calculated by:

$$Rate_{Cancer;\,COD} = \frac{D_{cancer;\,COD}}{Person\ Years}$$

The HR necessary to produce the same cancer rate estimated using the COD is:

$$Rate_{Cancer;\,COD} = Rate_{Cancer;\,PAF}$$

$$Rate_{Cancer;\,COD} = p_d \times \left(\frac{HR' - 1}{HR'}\right) \times Rate_{Overall}$$

$$\frac{1}{p_d}\frac{Rate_{Cancer;\,COD}}{Rate_{Overall}} = \frac{HR' - 1}{HR'}$$

$$\frac{Rate_{Cancer;\,COD}}{p_d Rate_{Overall}} = 1 - \frac{1}{HR'}$$

$$1 - \frac{Rate_{Cancer;\,COD}}{p_d Rate_{Overall}} = \frac{1}{HR'}$$

$$\frac{p_d Rate_{Overall} - Rate_{Cancer;\,COD}}{p_d Rate_{Overall}} = \frac{1}{HR'}$$

$$HR' = \frac{p_d Rate_{Overall}}{p_d Rate_{Overall} - Rate_{Cancer;\,COD}}$$

Overall, the cancer mortality rate calculated using COD from the cancer registries was 363.4 per 100,000 person-years compared to 519.7 per 100,000 person-years using the PAF. This results in cancer mortality rates about 30% lower than the cancer-attributable mortality calculated from the PAF (Table 7). The difference was fairly consistent among population subgroups and cancer sites. HR' was found overall to be 2.25, that is the HR would need to be reduced from 4.84 to 2.25 for the rates to be equivalent. HR' ranged from 1.94 among lung recipients to 2.47 among liver recipients.

							Using COD from cancer registries		Using adjusted PAF				
	Cancer deaths (n)	Deaths (n)	Deaths among those with cancer (n)	Person-years	p_d	Overall mortality rate per 100,000 person-years	Cancer deaths (n)	Cancer mortality mate per 100,000 PY	HR*	PAF*	95% CI	Cancer-attributable mortality rate per 100,000 person-years	HR'
Overall	11254	36749	6144	938084	0.17	3917.5	3409	363.4	4.84	0.13	(0.13, 0.13)	519.7	2.25
Age													
0-17	317	1360	102	79261	0.08	1715.9	52	65.6	5.8	0.06	(0.06, 0.06)	106.5	2.04
18-34	779	3155	294	156589	0.09	2014.8	147	93.9	5.49	0.08	(0.08, 0.08)	153.5	2
35-49	2722	9647	1259	309493	0.13	3117	718	232	5.06	0.1	(0.10, 0.10)	326.4	2.33
50-64	5777	17362	3414	327122	0.2	5307.5	1940	593.1	4.79	0.16	(0.16, 0.16)	825.8	2.32
65+	1659	5225	1075	65620	0.21	7962.5	552	841.2	4.1	0.16	(0.16, 0.16)	1238.6	2.06
Sex													
Female	3764	13588	1938	373121	0.14	3641.7	1057	283.3	4.88	0.11	(0.11, 0.11)	412.9	2.2
Male	7490	23161	4206	564964	0.18	4099.6	2352	416.3	4.83	0.14	(0.14, 0.14)	590.2	2.27
Race													
NH white	8075	25360	4585	602555	0.18	4208.7	2536	420.9	4.78	0.14	(0.14, 0.14)	601.6	2.24
NH black	1623	5817	803	135845	0.14	4282.1	438	322.4	4.35	0.11	(0.11, 0.11)	455.1	2.2
Hispanic	1087	4041	511	142165	0.13	2842.5	293	206.1	5.79	0.1	(0.10, 0.10)	297.3	2.34
API	421	1320	215	52289	0.16	2524.4	126	241	7.61	0.14	(0.14, 0.14)	357.1	2.42
Other	48	211	30	5230	0.14	4034.4	16	305.9	6.63	0.12	(0.12, 0.13)	487.1	2.14
Organ transplanted													
Kidney	6010	17257	2871	552597	0.17	3122.9	1643	297.3	4.67	0.13	(0.13, 0.13)	408.2	2.34
Liver	2020	7239	1166	178978	0.16	4044.6	694	387.8	5.84	0.13	(0.13, 0.13)	540	2.47
Heart	1943	6098	1280	111846	0.21	5452.1	662	591.9	4.76	0.17	(0.17, 0.17)	903.8	2.07
Lung	760	4056	567	33761	0.14	12014	274	811.6	3.99	0.1	(0.10, 0.10)	1258.2	1.94
Other	521	2099	260	60903	0.12	3446.5	136	223.3	5.73	0.1	(0.10, 0.10)	352.4	2.1

Table 7: Comparison of mortality rates estimated using COD and PAF from registries who provided COD information

Table 7 continued

	Cancer deaths (n)	Deaths (n)	Deaths among those with cancer (n)	Person-years	p_d	Overall mortality rate per 100,000 person-years	Using COD from cancer registries: Cancer deaths (n)	Using COD from cancer registries: Cancer mortality mate per 100,000 PY	Using adjusted PAF: HR*	Using adjusted PAF: PAF*	Using adjusted PAF: 95% CI	Using adjusted PAF: Cancer-attributable mortality rate per 100,000 person-years	HR'
Years since transplant													
0-2	3041	10964	979	288517	0.09	3800.1	638	221.1	12	0.08	(0.08, 0.08)	310.9	2.87
2-5	5050	10175	1621	301831	0.16	3371.1	1033	342.2	6.37	0.13	(0.13, 0.13)	452.7	2.76
5-10	6019	10585	2264	249869	0.21	4236.2	1214	485.9	4.14	0.16	(0.16, 0.16)	687.3	2.16
10+	3341	5025	1280	97867	0.25	5134.5	524	535.4	2.63	0.16	(0.16, 0.16)	810	1.69
Cancer site													
Lung	1554	36749	1276	938084	0.03	3917.5	943	100.5	10.5	0.03	(0.03, 0.03)	123	3.83
NHL	1689	36749	908	938084	0.02	3917.5	479	51.1	4.66	0.02	(0.02, 0.02)	76	2.12
CRC	650	36749	381	938084	0.01	3917.5	196	20.9	3.64	0.01	(0.01, 0.01)	29.4	2.06
Breast (wo	584	13588	204	373121	0.02	3641.7	82	22	1.88	0.01	(0.01, 0.01)	25.6	1.67
Kidney	894	36749	302	938084	0.01	3917.5	127	13.5	2.42	0	(0.00, 0.00)	18.9	1.73
Melanoma	425	36749	188	938084	0.01	3917.5	98	10.4	2.25	0	(0.00, 0.00)	11.1	2.09
Liver	143	36749	108	938084	0	3917.5	80	8.5	12.8	0	(0.00, 0.00)	10.6	3.86
Thyroid	262	36749	50	938084	0	3917.5	7	0.7	1.25	0	(0.00, 0.00)	1.1	1.16
Prostate (men only)	1219	23161	366	564964	0.02	4099.6	50	8.9	0.94	--			
All other cancers	5026	36749	2718	938084	0.07	3917.5	1393	148.5	3.67	0.05	(0.05, 0.05)	210.9	2.05
Infectious etiology†	2544	36749	1389	938084	0.04	3917.5	757	80.7	4.92	0.03	(0.03, 0.03)	117.9	2.2
Non-Infectious etiology	8710	36749	4755	938084	0.13	3917.5	2652	282.7	4.33	0.1	(0.10, 0.10)	389.8	2.26

HR=hazard ratio; PAF=population-attributable fraction; CI=confidence interval

*Adjusted for age at transplant, sex, race, transplant year, transplanted organ.

HR' is what the HR would need to be in order for the cancer mortality rates from the PAF and COD method to be equal

-- Since the HR is less than 1, the PAF is negative.

In order to assess if there are confounders that could plausibly reduce the HR from 4.84 to 2.25, I first estimated the HR adjusting for several additional observed covariates: Body mass index (BMI) and infection with Epstein-Barr virus (Table 8). Neither of these factors changed the HR by an amount that would result in a meaningful change to the cancer-attributable mortality rates.

	HR	**HR_{BMI}****	**HR_{EBV}**
Overall	4.84	4.93	4.84
Lung	10.48		
NHL	4.66	4.77	4.64
CRC	3.64	3.63	
Breast (women only)	1.88	1.88	
Kidney	2.42	2.48	
Melanoma	2.25		
Liver	12.82	11.57	
Thyroid	1.25		
Prostate	0.94		
Infectious etiology†	4.92		
Non-Infectious etiology	4.33		

HR=hazard ratio, BMI=body mass index, EV=Epstein-Barr virus
*All models adjusted for age at transplant, sex, race, transplant year, and organ transplanted.
**Cancers selected were those identified by the Annual Report to the Nation (2012) and the American Cancer Society as being associated with overweight and obesity
HR_{BMI} adjusted for BMI (underweight, normal,, overweight, obese)
HR_{EBV} adjusted for EBV status (positive, negative, unknown)

Table 8: Hazard ratio of death adjusted for potential confounders by cancer type

Second, unmeasured confounders could substantially reduce the HR. One method is to estimate the E-value, which is used to evaluate the potential impact of unmeasured confounders.[71] The E-value is the minimum association an unmeasured confounder must have with both cancer and mortality to diminish the association between cancer and mortality. A benefit of this approach is that it does not assume any structure of unmeasured confounders. The E-value is calculated as follows:

$$RR \approx \left(1 - 0.5^{\sqrt{HR}}\right) / \left(1 - 0.5^{\sqrt{1/HR}}\right)$$
$$E - value = RR + \sqrt{RR * (RR - 1)}$$

Using the HR of 4.14 from Table 7 results in an E-value of 4.68. To assess if there are potential confounders that can explain the association ask: "Is the association between a potential confounder and all cancers combined as strong as 4.68?" and "Is the association between a potential confounder and death as strong as 4.68?" The answer to both of these questions is no, therefore no plausible confounders exist with the magnitude required to explain the association between cancer and death.

A primary concern is the potential impact of smoking on lung cancer mortality. An extensive bias factor approach was undertaken in Chapter 4 and it showed that the unmeasured effects of smoking are unlikely to explain the difference in the methods or be a large source of bias.

In conclusion, I found a difference in the cancer-attributable mortality rates calculated using the PAF method and COD. One possibility is that the HR used in the PAF method was biased by unmeasured confounding. Two approaches were taken to determine if such a confounder is plausible. First, I found that a confounder would need to exist strong enough to reduce the HR from 4.84 to 2.25 in order for the cancer-attributable rates from both methods to be equal. Second, I used the E-value to determine that a potential confounder would need to have an association with cancer and death of at least 4.68. Both of these approaches indicated that a confounder would need a very strong association with both cancer and death to result in a large bias. It does not seem likely that any such confounder exists.

B. Coding of cause of death

The coding of COD may also contribute to a discrepancy in the cancer mortality rates estimated using COD compared to using the PAF. CODs from both recipients who developed cancer and those who did not were needed to investigate the coding of COD. These were available from the SRTR data for all registries and I compared them to the cancer registry CODs for the subset of 12 cancer registries who release COD information.

1) Comparison of cause of death from SRTR and cancer registries

There were 682 recipients with cancer coded as a COD in the SRTR but not recorded as having a cancer diagnosis from the cancer registries (Table 9). This could occur because the SRTR is recording PTLD or squamous cell carcinomas of the skin as cancer COD. Neither PTLD nor non-melanoma skin cancers are recorded as cancers by cancer registries and therefore classification as a cancer COD by the SRTR in the absence of incident cancer could reflect differing coding rules. Another possibility is that the recipient may have out-migrated from the cancer registry catchment area. This could result in the cancer registry not recording that incident cancer. Finally, the SRTR CODs are entered by clerks who may not have adequate expertise to assign COD and therefore may be recorded incorrectly.

COD from SRTR	No incident cancer		Incident cancer		Total
Cancer	682	20%	2773	80%	3455
Non-Cancer	22220	91%	2170	9%	15026
Missing (died but no COD)	7703	87%	1201	13%	8904
Total deaths	30605		6144		36749

Table 9: Distribution of COD from SRTR among those with incident cancer and those without

Table 10 shows the comparison of CODs from SRTR and cancer registries among those with an incident cancer. The modest kappa statistic (0.4) indicates that agreement between SRTR and cancer registries is only moderate.

	COD from cancer registries				
COD from SRTR	Non cancer		Cancer		Total
Non cancer	1425	67%	745	27%	2170
Cancer	711	33%	2062	73%	2773
Total	2136		2807		4943
Kappa=0.40					

Table 10: Comparison of COD from SRTR to COD from cancer registries among those with cancer

2) Fraction of deaths from other causes attributable to cancer

Some deaths listed in the SRTR as other causes may be, in fact, due to cancer. Using the PAF methodology, the proportion of deaths attributable to cancer was estimated for 5 mortality outcomes. This is discussed in more detail in Chapter 4. The results here are restricted to recipients from the 12 registries that reported COD.

A large proportion of deaths with a COD of cancer are attributed to cancer (79.7%; Table 11). This suggests that most, but not all deaths, with an SRTR COD of cancer are actually caused by the cancer. The proportion of deaths with an SRTR COD other than cancer that are attributed to cancer ranges from 9.2% of deaths with a missing cause to 4.3% of deaths with graft failure as the COD and 4.6% of deaths with infection as the COD. For example, deaths with COD of graft failure could occur as a result of reduction in immunosuppression in response to the cancer diagnosis, and deaths with COD of infection could be due to infections caused by cancer treatment. This analysis shows that there are deaths for which the underlying cause was likely cancer but the COD was coded as something else. Together, these other attributable deaths account for non-negligible mortality (24+20.4+90.6+87.6 = 222.6 per 100,000 person-years).

Mortality outcome category defined by COD in SRTR	**Deaths (n)**	**Number with cancer and died**	**Deaths with prior cancer, p_d (%)**	**HR for cancer**	**PAF (%)**	**Outcome-specific mortality rate**	**Cancer-attributable mortality rate**
Cancer	3455	2773	80.3	144.9	79.7	368.3	293.6
Infection	4910	379	7.7	2.5	4.6	523.4	24.0
Graft failure	4454	329	7.4	2.4	4.3	474.8	20.4
Other causes	15026	1462	9.7	2.4	5.7	1601.8	90.6
Missing	8904	1201	13.5	3.2	9.2	949.2	87.6
Overall	36749						516.1

HR=Hazard ratio; PAF=Population-attributable fraction
*Hazard ratios are computed using time since transplant as the time scale and are adjusted for sex, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian/Pacific Islander, other), organ transplanted (kidney, liver, heart, lung, other), calendar year of transplant (1987-1994, 1995-1999, 2000-2004, 2005-2008, 2009-2014) and age at transplant. Since the effect of age on mortality was non-linear, age was modeled as a piecewise linear function with knots at age 5 and 19 years.
Overall and cancer-attributable mortality rates are per 100,000 person-years

Table 11: Population attributable fraction and cancer-attributable mortality for other causes of death among 165,839 recipients with COD from cancer registries

Finally, the sum of all the cancer-attributable mortality rates across the CODs results in a mortality rate of 516.1 per 100,000 person-years. This is very close to the cancer-attributable mortality rate using any cause of death and the PAF method (519.7 per 100,000 person-years). This suggests that the difference in mortality rates from the PAF method compared to the COD method may be due to misattribution of cancer deaths into other COD codes.

C. Discussion

The difference between the cancer-attributable mortality rate derived from the PAF and the cancer-specific mortality estimated using the COD was investigated by understanding potential bias in the PAF and potential miscoding of COD. Based on several different approaches, it does not seem likely that the HR used in the PAF method is greatly biased by unmeasured confounding. There is, however, evidence to suggest that COD is incorrectly coded. The moderate agreement in COD from SRTR vs. cancer registries demonstrates that coding COD is challenging. Furthermore, many deaths can be attributed to cancer that were coded as other causes. The cancer-attributable mortality rate is higher than the cancer-specific mortality rate using COD because it correctly includes deaths which were attributable to cancer.

6. Life-years lost to cancer among solid organ transplant recipients[2]

A. Abstract

Background: Solid organ transplant recipients have an elevated risk of cancer. Quantifying the life-years lost (LYL) due to cancer provides a complementary view of the burden of cancer distinct from other metrics and may identify subgroups of transplant recipients who are most impacted.

Methods: Linked transplant and cancer registry data were used to identify incident cancers and deaths among solid organ transplant recipients in the United States (1987-2014). The LYL to cancer within 10 years after transplant were derived from the mean survival estimated from Cox models.

Results: Among 221,962 transplant recipients, 13,074 (5.9%) developed cancer within 10 years of transplant. During this period, the mean LYL was 0.16 years per transplant recipient and 2.7 years per cancer. Cancer accounted for 1.9% of the total LYL expected in this population. Lung cancer was the largest contributor, accounting for 24% of all LYL, and non-Hodgkin lymphoma had the next highest contribution (15%). Lung recipients had the highest proportion of total LYL to cancer, followed by heart recipients (4.4% and 2.8%, respectively). For all organ recipients, LYL to cancer increased with age. The youngest, age 0-34, had 0.5% of total LYL to cancer compared to 3.2% among those age 50+.

Conclusions: Transplant recipients have a shortened lifespan after developing cancer. Lung cancer and NHL are large contributors to the LYL to cancer, highlighting opportunities to reduce cancer mortality through prevention and screening.

---

[2] For submission to *Cancer*

B. Introduction

Solid organ transplant recipients have increased cancer risk compared with the general population.[1,81] Excess risk varies by cancer type and is largely due to immunosuppression from medications used to prevent rejection, underlying medical comorbidities, and end-stage organ disease.[15,16,19] Moreover, the solid organ transplant population in the United States has doubled since the late 1980s due to an increasing number of transplants and improved survival.[12] During 1987-1999 the cumulative incidence of cancer in the first 5 years after transplantation was 4.2%, and this rose slightly to 4.4% during 2000-2008.[82]

An additional important measure of the impact of cancer on the transplant population is cancer mortality since it reflects the final downstream outcome accounting for cancer incidence, treatment, and survival following cancer diagnosis. Among US transplant recipients during 1987-2014, 13% of deaths were attributed to cancer, corresponding to a cancer-attributable mortality rate of 516 per 100,000 person-years.[83] A relatively high proportion of deaths was specifically attributable to lung cancer (3.1%) and non-Hodgkin's lymphoma (NHL, 1.9%). Cancer-attributable mortality increased with age and time since transplant.

The benefit of transplantation among patients with end-stage organ disease can be quantified in terms of a gain in life-expectancy (i.e., life-years).[38] In a similar way, because recipients who develop cancer will have shorter lifetimes than they would have had without cancer, the adverse effect of cancer on mortality can be quantified by the life-years lost (LYL) to cancer. The LYL due to cancer in the population is the difference between the average actual lifetime with cancer and what would have been experienced in the population without cancer.[38,26] In other words, LYL estimates the extent to which the average lifespan is shortened by cancer.

In this study, we quantified the LYL due to cancer among US transplant recipients over the first 10 years post-transplant. We estimated the LYL by subgroups, including transplanted organ, age at transplant, sex, race/ethnicity, and cancer site. Evaluating the LYL due to cancer provides insight into which cancers are causing premature deaths and can help identify subgroups of transplant recipients in whom improved cancer prevention or treatment interventions may be most impactful.

C. Methods

The Transplant Cancer Match (TCM) Study links the US Scientific Registry of Transplant Recipients (SRTR) with state and regional cancer registries.[1] SRTR data include recipient demographic and transplant characteristics as well as follow-up for graft and overall survival. For the present analysis, 17 cancer registries provided data on incident cancers (excluding non-melanoma skin cancer) covering approximately half of the US transplant population for 1987-2014 (see Table 12 footnote). The TCM Study is considered non-human subjects research at the National Cancer Institute and was approved, as required, by participating cancer registries.

First primary invasive cancers in transplant recipients were identified from the linked cancer registries. Cancers were classified based on the primary site and histology of the tumor using a modified version of the Surveillance, Epidemiology and End Results site recode.[51] The study cohort comprised recipients who resided in areas covered by participating cancer registries at the time of transplant, received their first transplant during a period of cancer registry coverage, and did not previously have an invasive cancer recorded in the cancer registry. Prior cancers were those diagnosed before transplant or within 90 days post-transplant (since those likely developed pre-transplantation). Follow-up thus started 90 days post-transplant, and ended at the earliest of death, loss to follow-up by the SRTR, end of cancer registry coverage, or 10 years post-transplant.

For this cohort of transplant recipients, we calculated the LYL due to cancer as the difference in average survival between a hypothetical scenario in which recipients were cancer-free, and the actual scenario in which some recipients developed cancer. For the cancer-free scenario, follow-up for recipients who developed cancer was censored at the time of their cancer diagnosis. For each scenario, the mean survival was estimated using Cox proportional hazards regression models. Both models were calculated separately according to transplanted organ (kidney, liver, heart, lung, other/multiple), and the covariates were sex, race/ethnicity, year of transplant (1987-94, 1995-99, 2000-04, 2005-08, 2009-14), body mass index (BMI; <18.5, 18.5-24.9, 25-29.9, 30+), diabetes status, cancer registry, and organ-specific characteristics and reasons for transplant (Table 12 and Table 13).

	Total		Kidney		Liver		Heart		Lung		Other/ multiple	
	(n=221,962)		(n=136,534)		(n=391,55)		(n=22,912)		(n=10,467)		(n=12,894)	
Sex												
Male	135,228	61%	81,563	60%	23,278	59%	17,253	75%	5571	53%	7563	59%
Female	86,734	39%	54,971	40%	15,877	41%	5659	25%	4896	47%	5331	41%
Age at transplant, years												
0-17	16,707	8%	7917	6%	4632	12%	3045	13%	402	4%	711	6%
18-34	34,997	16%	25,599	19%	2816	7%	2003	9%	1345	13%	3234	25%
35-49	67,595	30%	42,826	31%	11,923	30%	4676	20%	2021	19%	6149	48%
50-64	82,092	37%	46,064	34%	16,933	43%	11,029	48%	5489	52%	2577	20%
65-96	20,571	9%	14,128	10%	2851	7%	2159	9%	1210	12%	223	2%
Race/ethnicity												
Non-Hispanic White	134,535	61%	71,751	53%	27,794	71%	16,574	72%	8907	85%	9509	74%
Non-Hispanic Black	40,114	18%	31,021	23%	3366	9%	3433	15%	750	7%	1544	12%
Hispanic	33,726	15%	23,549	17%	6085	16%	2064	9%	611	6%	1417	11%
Asian/Pacific Islander	12,146	5%	9374	7%	1631	4%	718	3%	149	1%	274	2%
Other	1,441	1%	839	1%	279	1%	123	1%	50	0%	150	1%
Transplant year												
1987-1994	32,257	15%	19,902	15%	5774	15%	4285	19%	903	9%	1393	11%
1995-1999	49,140	22%	29,457	22%	8997	23%	5691	25%	1941	19%	3054	24%
2000-2004	58,297	26%	35,729	26%	10,897	28%	5624	25%	2539	24%	3508	27%
2005-2008	50,901	23%	31,809	23%	8715	22%	4443	19%	2922	28%	3012	23%
2009-2014	31,367	14%	19,637	14%	4772	12%	2869	13%	2162	21%	1927	15%

Table 12: Characteristics of transplant recipients in the United States, by transplanted organ

Table 12 continued

	Total (n=221,962)		Kidney (n=136,534)		Liver (n=391,55)		Heart (n=22,912)		Lung (n=10,467)		Other/ multiple (n=12,894)	
Body mass index, kg/m^2												
Underweight (<18.5)	8,246	4%	3821	3%	1200	3%	1353	6%	1341	13%	531	4%
Normal (18.5-24.9)	77,170	35%	44,611	33%	12,940	33%	8933	39%	4500	43%	6186	48%
Overweight (25-29.9)	60,417	27%	35,627	26%	11,455	29%	6990	31%	3087	29%	3258	25%
Obese (30+)	44,004	20%	28,184	21%	9695	25%	3663	16%	1217	12%	1245	10%
Not recorded	32,125	14%	24,291	18%	3865	10%	1973	9%	322	3%	1674	13%
Diabetes mellitus												
Yes	50,364	23%	31,359	23%	5372	14%	3370	15%	1183	11%	9080	70%
No	131,975	59%	80,003	59%	26,826	69%	14,669	64%	8180	78%	2297	18%
Missing	7,366	3%	5270	4%	1183	3%	588	3%	201	2%	124	1%
Unknown (not recorded before 1995)	32,257	15%	19902	15%	5774	15%	4285	19%	903	9%	1393	11%

This study includes data from 17 US population-based cancer registries: California (years of cancer registry data 1988-2012), Colorado (1988-2009), Connecticut (1973- 2009), Florida (1981-2009), Georgia (1995-2010), Hawaii (1973-2007), Illinois (1986-2013), Iowa (1973-2009), Kentucky (1995-2011), Michigan (1985-2009), New Jersey (1979-2010), New York (1976- 2010), North Carolina (1990-2010), Pennsylvania (1985-2013), Texas (1995-2010), Utah (1973-2008), and the Seattle-Puget Sound area of Washington (1974- 2014).

Table 13: Additional characteristics of transplant recipients by organ transplanted

	Kidney (n=136,534)		Liver (n=39,155)		Heart (n=22,912)	Lung (n=10,467)	Other/multi (n=12,894)
Donor type							
Deceased	86,633	63%					
Living	49,901	37%					
Years from initial dialysis to transplant							
No dialysis	15,292	5%					
0.01-0.99	28,465	10%					
1-1.99	26,375	11%					
2-2.99	18,796	9%					
3-3.99	13,603	7%					
4-3.99	9790	5%					
5+	22,766	13%					
Unknown	1447	1%					
Reason for transplant							
Glomerular diseases	40,504	30%					
Diabetes	32,193	24%					
Polycystic kidneys	12,907	9%					
Hypertensive nephrosclerosis	26,969	20%					
Other	19,580	14%					
Viral hepatitis			13,474	34%			
Alcohol use			10,128	26%			
PBC/PSC			4838	12%			
Other			15,625	40%			

Table 13 continued

	Kidney (n=136,534)		Liver (n=39,155)		Heart (n=22,912)		Lung (n=10,467)		Other/multi (n=12,894)	
Cardiomyopathy					12,193	53%				
Coronary artery disease					10,001	44%				
Other					1899	8%				
COPD							3817	36%		
Cystic fibrosis							1584	15%		
Idiopathic pulmonary fibrosis							2892	28%		
Other							2633	25%		
Tobacco use										
Yes							5442	52%		
No							2937	28%		
Unknown							2088	20%		
Lung transplant										
Single							4571	44%		
Double							5896	56%		

This study includes data from 17 US population-based cancer registries: California (years of cancer registry data 1988-2012), Colorado (1988-2009), Connecticut (1973- 2009), Florida (1981-2009), Georgia (1995-2010), Hawaii (1973-2007), Illinois (1986- 2013), Iowa (1973-2009), Kentucky (1995-2011), Michigan (1985-2009), New Jersey (1979-2010), New York (1976- 2010), North Carolina (1990-2010), Pennsylvania (1985-2013), Texas (1995-2010), Utah (1973-2008), and the Seattle-Puget Sound area of Washington State (1974- 2014)

The models were also adjusted for age at transplant, which was modeled as a piecewise linear function with knots at age 5 and 19. The time scale for the Cox models was time since transplant. The hazard ratios from the Cox models estimating the overall and cancer-free models are shown in Table 14 to Table 16.

For each model, the direct adjusted survival curve was estimated and averaged over the individual recipient profiles.[84] The restricted mean survival time for each curve was then estimated as the area under the curve up to 10 years post-transplant.[85] We censored the follow-up time at 10 years to avoid estimating the tail of the survival curves where data were sparse. Finally, the mean LYL per transplant recipient was derived as the difference in the areas under the curves between the hypothetical cancer-free and actual scenarios.

Multiple imputation was used to handle missing data on recipient characteristics shown in Table 12 and Table 13. BMI had a moderate amount of missing values ranging from only 3% among lung recipients to 18% among kidney recipients. Tobacco use was most frequently missing with 20% unknown for lung recipients. Five organ-specific imputed datasets were generated using the sequential regression imputation method implemented with IVEware.[86] The imputation models included the organ-specific covariates, the censoring indicator, and time since transplant. The analysis described above was conducted for each imputed dataset separately, and the within-imputation variance was estimated using 250 bootstrap samples. The results were then combined using Rubin's rule in SAS/PROC MIANALYZE.[53]

	Kidney				Liver			
	Overall survival		Cancer-free survival		Overall survival		Cancer-free survival	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Sex (ref: Male)								
Female	0.9	0.88, 0.93	0.93	0.90, 0.96	1	0.95, 1.04	1.01	0.96, 1.06
Race (ref: Non-Hispanic white)								
Non-Hispanic black	0.94	0.91, 0.98	0.94	0.90, 0.97	1.38	1.28, 1.48	1.44	1.34, 1.56
Hispanic	0.68	0.65, 0.71	0.68	0.64, 0.71	0.91	0.85, 0.97	0.93	0.87, 1.00
Asian/Pacific Islander	0.6	0.56, 0.65	0.59	0.55, 0.64	0.7	0.62, 0.79	0.69	0.61, 0.79
Other	0.94	0.80, 1.11	0.92	0.76, 1.10	0.96	0.74, 1.25	0.98	0.74, 1.29
Age (1 year increase)								
Age ≤ 5	0.94	0.84, 1.05	0.92	0.82, 1.04	1	0.94, 1.07	1	0.93, 1.06
Age 5-18	1.08	0.95, 1.22	1.09	0.96, 1.25	1.05	0.97, 1.13	1.07	0.98, 1.16
Age >18	1.04	1.02, 1.06	1.04	1.02, 1.06	0.97	0.95, 0.99	0.96	0.94, 0.98
Transplant year (ref: 1987-1994)								
1995-1999	0.58	0.12, 2.77	0.59	0.12, 2.93	0.86	0.81, 0.91	0.84	0.79, 0.90
2000-2004	0.47	0.10, 2.30	0.47	0.09, 2.41	0.78	0.74, 0.83	0.77	0.71, 0.82
2005-2008	0.35	0.07, 1.71	0.34	0.07, 1.72	0.82	0.76, 0.88	0.82	0.75, 0.88
2009-2014	0.25	0.05, 1.23	0.24	0.05, 1.23	0.69	0.61, 0.78	0.68	0.60, 0.77
BMI (ref=underweight)								
Normal	0.8	0.72, 0.90	0.8	0.72, 0.88	0.73	0.64, 0.82	0.75	0.65, 0.86
Overweight	0.77	0.69, 0.85	0.77	0.70, 0.85	0.65	0.58, 0.73	0.68	0.59, 0.78
Obese	0.84	0.75, 0.93	0.85	0.77, 0.94	0.67	0.59, 0.76	0.7	0.61, 0.80
Diabetes (ref=No)								
Unknown (after 1995)	0.75	0.15, 3.73	0.77	0.15, 4.04				
Yes	1.13	1.07, 1.20	1.17	1.10, 1.24	1.41	1.33, 1.49	1.49	1.40, 1.59
Registry (ref: Utah)								
California	1.03	0.90, 1.17	1	0.87, 1.16	1.03	0.83, 1.29	1.11	0.86, 1.42
Colorado	1.07	0.92, 1.26	1.03	0.87, 1.22	0.92	0.72, 1.20	1.02	0.76, 1.36
Connecticut	1.09	0.93, 1.28	1.08	0.91, 1.29	1.22	0.93, 1.59	1.34	0.99, 1.81
Florida	1.16	1.02, 1.33	1.11	0.96, 1.28	1.11	0.89, 1.40	1.2	0.93, 1.55
Georgia	1.1	0.95, 1.28	1.08	0.91, 1.26	1.33	1.04, 1.70	1.49	1.13, 1.96
Hawaii	1.45	1.15, 1.82	1.43	1.11, 1.83	1.33	0.89, 1.98	1.48	0.96, 2.29
Iowa	0.97	0.83, 1.14	0.95	0.80, 1.13	1.11	0.85, 1.43	1.19	0.88, 1.59
Illinois	1.14	1.00, 1.30	1.07	0.93, 1.24	1.14	0.91, 1.43	1.18	0.91, 1.53
Kentucky	1.49	1.28, 1.74	1.46	1.23, 1.72	1.09	0.84, 1.40	1.16	0.87, 1.55
Michigan	1.2	1.05, 1.37	1.14	0.99, 1.33	1.19	0.94, 1.50	1.24	0.95, 1.61
North Carolina	1.03	0.89, 1.19	1	0.85, 1.17	1	0.79, 1.27	1.09	0.83, 1.43

Table 14: Hazard ratios from the Cox proportional hazards models used to estimate the overall survival and cancer-free survival experience for kidney and liver recipients

Table 14 continued

	Kidney				Liver			
	Overall survival		Cancer-free survival		Overall survival		Cancer-free survival	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
New Jersey	1.21	1.06, 1.40	1.17	1.01, 1.37	1.16	0.92, 1.47	1.24	0.95, 1.61
New York	1.02	0.90, 1.17	1	0.86, 1.15	1.22	0.98, 1.52	1.31	1.02, 1.69
Pennsylvania	1.16	1.01, 1.32	1.12	0.97, 1.29	1.28	1.02, 1.59	1.38	1.07, 1.77
Seattle	0.99	0.85, 1.16	0.93	0.79, 1.10	1.14	0.89, 1.46	1.14	0.86, 1.51
Texas	1.11	0.97, 1.28	1.09	0.94, 1.27	1.22	0.98, 1.54	1.31	1.01, 1.69
Donor type (ref=Living)	1.28	1.24, 1.33	1.34	1.29, 1.39				
Years from initial dialysis to transplant (ref=No dialysis)								
0.01-0.99	1.5	1.40, 1.60	1.51	1.40, 1.63				
1-1.99	1.79	1.67, 1.91	1.84	1.71, 1.98				
2-2.99	1.92	1.79, 2.06	2.01	1.86, 2.17				
3-3.99	2.1	1.95, 2.26	2.22	2.04, 2.41				
4-3.99	2.22	2.05, 2.40	2.39	2.19, 2.61				
5+	2.69	2.50, 2.90	2.9	2.67, 3.14				
Reasons for Transplant (ref=No)								
Glomerular diseases	0.77	0.73, 0.80	0.76	0.72, 0.80				
Diabetes	1.67	1.58, 1.76	1.79	1.69, 1.89				
Polycystic kidneys	0.64	0.61, 0.68	0.62	0.58, 0.66				
Hypertensive nephrosclerosis	1.02	0.98, 1.06	1.05	1.00, 1.10				
Other	0.89	0.85, 0.94	0.89	0.85, 0.94				
Viral hepatitis					1.26	1.19, 1.33	1.33	1.25, 1.41
Alcohol use					1.2	1.13, 1.27	1.13	1.06, 1.21
PBC/PSC					0.69	0.64, 0.75	0.65	0.59, 0.71
Other					0.9	0.85, 0.95	0.91	0.85, 0.97

Abbreviations: PBC=primary biliary cirrhosis, PSC=primary sclerosing cholangitis

Cox models used time since transplant as the time scale and were adjusted for sex, race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian/Pacific Islander, other), calendar year of transplant ((1987-1994, 1995-1999, 2000-2004, 2005-2008, 2009-2014), cancer registry, organ-specific covariates, and age at transplant. Since the effect of age on mortality was non-linear, age was modeled as a piecewise linear function with knots at age 5 and 19 years.

	Heart				Lung			
	Overall survival		Cancer-free survival		Overall survival		Cancer-free survival	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Sex (ref: Male)								
Female	1.03	0.98, 1.09	1.07	1.01, 1.14	0.97	0.91, 1.02	1.02	0.90, 1.15
Race (ref: Non-Hispanic white)								
Non-Hispanic black	1.63	1.53, 1.73	1.75	1.64, 1.87	1.02	0.91, 1.15	1.08	0.94, 1.25
Hispanic	1.09	0.99, 1.19	1.12	1.01, 1.23	1.06	0.92, 1.21	0.87	0.65, 1.18
Asian/Pacific Islander	1	0.86, 1.16	1.05	0.89, 1.23	0.96	0.73, 1.25	1.32	0.86, 2.01
Other	1.11	0.80, 1.54	1.2	0.85, 1.71	1.41	0.96, 2.08	1.11	0.99, 1.25
Age (1 year increase)								
Age ≤ 5	1.03	0.99, 1.08	1.02	0.98, 1.07	1.13	1.01, 1.27	0.87	0.76, 0.99
Age 5-18	0.98	0.93, 1.04	1	0.94, 1.06	0.85	0.75, 0.97	1.04	1.01, 1.07
Age >18	0.99	0.98, 1.01	0.98	0.96, 0.99	1.05	1.02, 1.08	0.96	0.10, 9.47
Transplant year (ref: 1987-1994)								
1995-1999	0.82	0.77, 0.87	0.83	0.78, 0.89	0.96	0.10, 8.86	0.73	0.07, 7.66
2000-2004	0.75	0.70, 0.80	0.78	0.72, 0.84	0.77	0.08, 7.53	0.74	0.07, 7.70
2005-2008	0.66	0.60, 0.72	0.67	0.61, 0.74	0.75	0.08, 7.34	0.67	0.06, 6.98
2009-2014	0.54	0.46, 0.63	0.56	0.48, 0.66	0.68	0.07, 6.64	0.9	0.81, 0.99
BMI (ref=underweight)								
Normal	0.95	0.85, 1.05	0.95	0.85, 1.06	0.88	0.80, 0.97	0.97	0.87, 1.09
Overweight	1.01	0.90, 1.12	1.03	0.92, 1.16	0.95	0.86, 1.06	1.01	0.88, 1.17
Obese	1.07	0.96, 1.21	1.1	0.98, 1.25	0.98	0.86, 1.12	1.16	1.03, 1.29
Diabetes (ref=No)								
Yes	1	1.00, 1.00	1.28	1.19, 1.39	1.17	1.06, 1.30	1.22	0.82, 1.81
Registry (ref: Utah)								
California	1.12	0.92, 1.37	1.06	0.85, 1.32	1.26	0.91, 1.75	1.39	0.96, 2.00
Colorado	1.03	0.81, 1.32	1	0.76, 1.32	1.14	0.80, 1.63	1.22	0.82, 1.81
Connecticut	1.16	0.90, 1.48	1.16	0.88, 1.52	1.23	0.82, 1.85	1.39	0.89, 2.17
Florida	0.99	0.80, 1.21	0.94	0.75, 1.19	0.92	0.66, 1.29	0.97	0.66, 1.42
Georgia	1.08	0.85, 1.37	1.03	0.79, 1.35	1.25	0.87, 1.78	1.38	0.93, 2.05
Hawaii	1.74	1.12, 2.70	1.5	0.91, 2.48	2.34	0.91, 6.07	2.83	1.07, 7.48
Iowa	1.01	0.78, 1.31	0.91	0.68, 1.22	0.99	0.67, 1.47	1.06	0.68, 1.64
Illinois	1.18	0.96, 1.45	1.13	0.90, 1.43	0.98	0.71, 1.37	1.06	0.73, 1.54
Kentucky	1.44	1.13, 1.83	1.44	1.10, 1.88	1.47	1.04, 2.08	1.59	1.08, 2.35
Michigan	1.11	0.89, 1.38	1.04	0.81, 1.33	1	0.71, 1.41	1.04	0.71, 1.52
North Carolina	1.14	0.92, 1.41	1.07	0.84, 1.36	1.15	0.82, 1.61	1.24	0.84, 1.81
New Jersey	1	0.81, 1.24	0.93	0.74, 1.19	1.12	0.79, 1.59	1.21	0.82, 1.80
New York	1.11	0.90, 1.36	1.08	0.86, 1.36	1.23	0.88, 1.72	1.32	0.90, 1.92
Pennsylvania	1.11	0.90, 1.36	1.05	0.84, 1.32	1.11	0.80, 1.54	1.19	0.82, 1.73
Seattle	0.66	0.51, 0.85	0.51	0.38, 0.69	1.02	0.72, 1.45	1.07	0.72, 1.58
Texas	1.24	1.01, 1.52	1.24	0.99, 1.57	1.33	0.96, 1.85	1.42	0.98, 2.06

Table 15: Hazard ratios from the Cox proportional hazards models used to estimate the overall survival and cancer-free survival experience for heart and lung recipients

Table 15 continued

	Heart				Lung			
	Overall survival		Cancer-free survival		Overall survival		Cancer-free survival	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Cardiomyopathy	0.9	0.83, 0.97	0.91	0.84, 0.99				
Coronary artery disease	1.09	1.01, 1.18	1.1	1.01, 1.20				
Other	0.89	0.80, 0.98	0.92	0.82, 1.03				
Tobacco use (ref=Yes)					0.99	0.91, 1.08	1.02	0.93, 1.11
Lung transplant (ref=double)					1.31	1.23, 1.41	1.26	1.17, 1.36
Reasons for transplant								
COPD (ref=No)					1.14	1.01, 1.27	1.14	1.01, 1.29
Cystic fibrosis					1.15	0.99, 1.34	1.11	0.94, 1.30
Idiopathic pulmonary fibrosis (ref=No)					1.11	0.99, 1.25	1.07	0.94, 1.21
Other (ref=No)					1.01	0.91, 1.12	1	0.90, 1.12

Cox models used time since transplant as the time scale and were adjusted for sex, race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian/Pacific Islander, other), calendar year of transplant ((1987-1994, 1995-1999, 2000-2004, 2005-2008, 2009-2014), cancer registry, organ-specific covariates, and age at transplant. Since the effect of age on mortality was non-linear, age was modeled as a piecewise linear function with knots at age 5 and 19 years.

		Other/multiple organs			
		Overall survival		Cancer-free survival	
		HR	95% CI	HR	95% CI
Sex (ref: Male)	Female	1.04	0.96, 1.12	1.04	0.96, 1.14
Race (ref: Non-Hispanic white)	Non-Hispanic black	1.14	1.00, 1.29	1.15	1.01, 1.32
	Hispanic	0.95	0.83, 1.09	0.93	0.80, 1.08
	Asian/Pacific Islander	1.04	0.78, 1.38	1.13	0.84, 1.52
	Other	0.90	0.64, 1.28	0.95	0.66, 1.36
Age (1 year increase)	Age ≤ 5	1.00	0.92, 1.09	1.00	0.92, 1.09
	Age 5-18	0.94	0.85, 1.04	0.94	0.85, 1.04
	Age >18	1.08	1.05, 1.11	1.08	1.05, 1.11
Transplant year (ref: 1987-1994)	1995-1999	1.35	1.17, 1.56	1.36	1.16, 1.58
	2000-2004	1.10	0.94, 1.28	1.11	0.94, 1.30
	2005-2008	0.96	0.81, 1.14	0.97	0.81, 1.16
	2009-2014	0.89	0.71, 1.12	0.89	0.70, 1.14
BMI (ref=underweight)	Normal	0.68	0.57, 0.81	0.67	0.56, 0.80
	Overweight	0.65	0.55, 0.78	0.64	0.53, 0.77
	Obese	0.70	0.55, 0.88	0.72	0.56, 0.91
Diabetes (ref=No)	Unknown (after 1995)				
	Yes	0.54	0.49, 0.61	0.54	0.48, 0.61
Registry (ref: Utah)	California	1.09	0.74, 1.61	1.04	0.69, 1.58
	Colorado	0.81	0.50, 1.31	0.81	0.49, 1.35
	Connecticut	1.50	0.92, 2.45	1.42	0.84, 2.40
	Florida	1.21	0.81, 1.82	1.22	0.79, 1.87
	Georgia	0.95	0.61, 1.48	0.90	0.56, 1.44
	Hawaii	1.29	0.55, 3.03	0.80	0.27, 2.35
	Iowa	0.89	0.56, 1.42	0.84	0.51, 1.37
	Illinois	1.09	0.73, 1.61	1.06	0.70, 1.60
	Kentucky	1.56	1.00, 2.43	1.57	0.99, 2.51
	Michigan	1.12	0.74, 1.68	1.08	0.70, 1.68
	North Carolina	0.94	0.62, 1.43	0.93	0.59, 1.45
	New Jersey	1.10	0.72, 1.68	1.11	0.71, 1.73
	New York	1.20	0.80, 1.81	1.16	0.75, 1.78
	Pennsylvania	1.15	0.77, 1.70	1.07	0.71, 1.63
	Seattle	0.99	0.65, 1.52	0.92	0.58, 1.45
	Texas	1.04	0.69, 1.57	1.04	0.67, 1.61

Cox models used time since transplant as the time scale and were adjusted for sex, race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian/Pacific Islander, other), calendar year of transplant ((1987-1994, 1995-1999, 2000-2004, 2005-2008, 2009-2014), cancer registry, organ-specific covariates, and age at transplant. Since the effect of age on mortality was non-linear, age was modeled as a piecewise linear function with knots at age 5 and 19 years.

Table 16: Hazard ratios from the Cox proportional hazards models used to estimate the overall survival and cancer-free survival experience for recipients of other or multiple organs

We calculated the proportion of total LYL due to cancer by dividing the total LYL by the expected person-years in a cancer-free cohort. The expected person-years was calculated as the mean survival of the cancer-free cohort multiplied by the number of recipients. The estimates were produced by transplanted organ, age group, sex, race/ethnicity, and specific cancer sites (colorectum, liver, lung, melanoma, breast, prostate, kidney, thyroid, non-Hodgkin lymphoma [NHL], and cancers of infectious etiology [see Table 12 note]). We also estimated the LYL per cancer, i.e., total LYL divided by the number of incident cancers. Finally, we estimated a cancer site's contribution to the LYL to cancer (i.e. the proportion of the LYL to cancer) by dividing the LYL for that specific cancer site by the total LYL due to cancer. To obtain estimates that combined recipients of different organs, the LYL for recipients of each organ were summed.

We present a sensitivity analysis using a different method, based on incorporating Kaplan-Meier curves showing the survival estimates for transplant recipients under the actual scenario and the hypothetical cancer-free scenario (i.e., censoring at the first cancer diagnosis). Using these curves, we derived an alternative non-parametric estimate of mean LYL per transplant recipient as the difference in the areas under the curves up to 10 years post-transplant. The 95% confidence interval (95%CI) for the difference was estimated using 250 bootstrap samples. Computations were performed using SAS software version 9.4 (Cary, NC).

D. Results

This study included 221,962 solid organ transplant recipients (Table 12). The majority were male (61%) and non-Hispanic White (61%). The most frequently transplanted organs were kidney (62%), liver (18%), heart (10%) and lung (5%). Overall, about 30% of recipients were 35-49 years old and 46% were age 50 or older; while heart and lung recipients were older than recipients of other organs. Eight percent of recipients were children under age 18.

This cohort of transplant recipients was followed for 1,118,284 person-years within 10 years of transplant (mean 5.0 years per person). Overall, 13,074 recipients (5.9%) developed cancer during this period. In total, 35,371 life-years were lost to cancer. Thus, the mean LYL was 0.16 years per transplant recipient within this 10-year follow-up period (95%CI: 0.15, 0.17) and 2.7 years per cancer (Table 17). This resulted in a loss due to cancer of 1.9% of all life-years that would be expected under a cancer-free scenario.

	Transplant recipients, n	Cancer cases, n (% of cohort)	Total LYL	LYL per transplant recipient		Proportion of total life-years lost to cancer, %	
				Mean	**95% CI**	**Mean**	**95% CI**
Overall	221,962	13,074 (5.9%)	35,371	0.16	0.15, 0.17	1.9	1.8, 1.9
Transplanted organ							
Kidney	136,534	7127 (5.2%)	19,148	0.14	0.13, 0.15	1.6	1.5, 1.6
Liver	39,155	2366 (6.0%)	6890	0.18	0.16, 0.19	2.1	1.9, 2.2
Heart	22,912	2182 (9.5%)	5171	0.23	0.21, 0.24	2.8	2.5, 3.0
Lung	10,467	875 (8.4%)	2878	0.27	0.24, 0.31	4.4	3.9, 4.9
Other/multiple	12,894	524 (4.1%)	1445	0.11	0.09, 0.13	1.3	1.1, 1.5
Sex							
Male	135,228	8802 (6.5%)	24,647	0.18	0.17, 0.19	2.1	3.0, 2.2
Female	86,734	4272 (4.9%)	10,674	0.12	0.12, 0.13	1.3	1.3, 1.5

Table 17: Life-years lost to cancer among US solid organ transplant recipients

Table 17 continued

	Transplant recipients, n	Cancer cases, n (% of cohort)	Total LYL	LYL per transplant recipient		Proportion of total life-years lost to cancer, %	
				Mean	95% CI	Mean	95% CI
Age at transplant, years							
0-34	51,704	1186 (2.0%)	2226	0.04	0.04, 0.05	0.5	0.5, 0.6
35-49	67,595	2946 (4.0%)	6442	0.1	0.09, 0.10	1.1	1.0, 1.1
50+	102,663	8942 (4.0%)	26,837	0.26	0.25, 0.27	3.2	3.2, 3.3
Race/ethnicity							
Non-Hispanic White	134,535	9248 (6.9%)	24,750	0.18	0.18, 0.19	2.2	2.1, 2.2
Non-Hispanic Black	40,114	2047 (5.1%)	5334	0.13	0.13, 0.14	1.6	1.5, 1.7
Hispanic	33,726	1256 (3.7%)	3527	0.1	0.09, 0.12	1.2	1.0, 1.3
Asian/Pacific Islander and other	13,587	523 (3.8%)	1728	0.13	0.11, 0.15	1.4	1.2, 1.6

LYL=life-years lost; CI=confidence interval

Mean life-years lost were derived from organ-specific Cox models using time since transplant as the time scale and included variables for sex, race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian/Pacific Islander, other), calendar year of transplant (1987-1994, 1995-1999, 2000-2004, 2005-2008, 2009-2014), cancer registry, organ-specific covariates, and age at transplant. Since the effect of age on mortality was non-linear, age was modeled as a piecewise linear function with knots at age 5 and 19 years.

The sum of the LYL across subgroups does not equal the total because the LYL depends on non-linear contributions from the Cox models.

The Kaplan-Meier curves for the actual cohort experience and hypothetical cancer-free cohort are shown in Figure 8, where the LYL to cancer is the area between the two curves (0.19 years per transplant recipient).

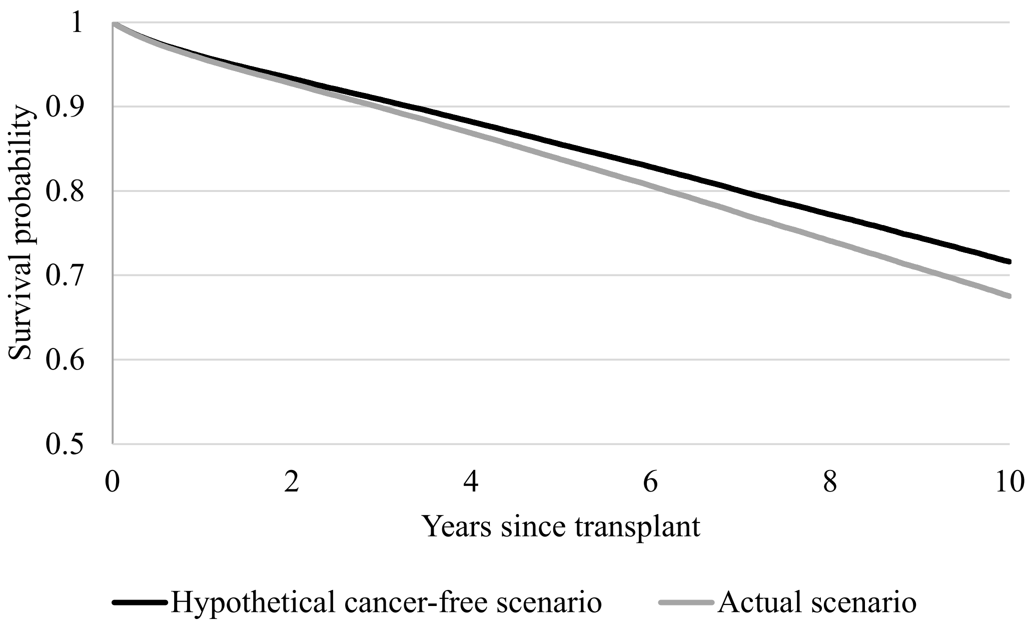


Figure 8: Kaplan-Meier curves for the cohort of transplant recipients under actual and hypothetical cancer-free scenarios

LYL varied by cancer site (Table 18). Lung cancer contributed the largest life-years lost in the cohort (0.45% of total expected life-years, or 24% of all LYL to cancer). NHL had the next highest contribution (0.29% of total expected life-years, or 15% of all LYL to cancer). Cancers of infectious etiology, which include NHL, were associated with 0.43% of total expected life-years (23% of all LYL to cancer). Among recipients who developed cancer, those with lung or liver cancer had the highest LYL to their cancer (4.7 and 4.9 years, respectively), while recipients who developed NHL, colorectal cancer, or melanoma lost 1.6-2.8 years (Table 18). Finally, recipients who developed kidney, thyroid, or breast cancer lost 0.5-1.3 life-years from their cancer. Prostate cancer was not statistically significantly associated with LYL.

Cancer site	Cancer cases, n (% of cohort)	Total LYL	LYL per transplant recipient		Proportion of total life-years lost to cancer, %		LYL per cancer case	
			Mean	95% CI	Mean	95% CI	Mean	95% CI
Lung	1788 (0.8%)	8401	0.035	0.035, 0.040	0.45	0.42, 0.48	4.7	4.5, 4.9
NHL	1970 (0.9%)	5430	0.022	0.022, 0.027	0.29	0.26, 0.32	2.8	2.5, 3.0
Colorectum	742 (0.3%)	1639	0.007	0.006, 0.009	0.09	0.07, 0.10	2.2	1.9, 2.5
Kidney	1018 (0.5%)	1348	0.006	0.005, 0.007	0.07	0.06, 0.09	1.3	1.1, 1.6
Melanoma	524 (0.2%)	830	0.004	0.003, 0.005	0.04	0.03, 0.06	1.6	1.2, 2.0
Liver	149 (0.1%)	732	0.003	0.003, 0.004	0.04	0.03, 0.05	4.9	4.1, 5.7
Breast	695 (0.3%)	571	0.003	0.002, 0.003	0.03	0.02, 0.04	0.8	0.6, 1.1
Thyroid	311 (0.1%)	160	0.0007	0.0002, 0.001	0.01	0.002, 0.02	0.5	0.1, 0.9
Prostate	1500 (0.7%)	46	0.0002	-0.001, 0.001	0	-0.01, 0.01	0.03	-0.1, 0.2
Infectious etiology*	2924 (1.3%)	8161	0.037	0.034, 0.040	0.43	0.40, 0.47	2.8	2.6, 3.0

LYL=life-years lost; CI=confidence interval, NHL=non-Hodgkin lymphoma

*Cancers with infectious etiology include NHL, Hodgkin lymphoma, Kaposi sarcoma, and cancers of the nasopharynx, liver, stomach, oropharynx including tonsil, anus, vulva, cervix, penis, and vagina.

Table 18: Life-years to cancer among US solid organ transplant recipients, by cancer site

The LYL results varied by transplanted organ (Table 17). Lung recipients had the highest mean LYL per person (0.27 years). This also corresponded to the highest proportion of life-years lost due to cancer (4.4% of total expected life-years). Furthermore, lung recipients comprised 5% of the entire cohort but disproportionately had 8.1% of the total LYL. The mean LYL per person was similar for heart recipients (0.23 years), and the proportion of life-years lost to cancer among heart recipients was 2.8%. However, since there were more heart than lung recipients, 15% of the total LYL among all organ recipients occurred among heart recipients. Kidney recipients had an average of 0.14 LYL per person and 1.6% of expected life-years were lost due to cancer. The results were very similar in our sensitivity analysis based on the difference in areas between Kaplan-Meier curves for the actual and hypothetical cancer-free scenarios (Table 19). Measures of LYL to cancer increased with age (Table 17). The youngest age group had a proportion of total expected LYL to cancer of 0.5%. This increased to 1.1% LYL to cancer among 35-49 age group, while those age 50 years and older had the greatest LYL to cancer (3.2% of expected life-years). In addition, the oldest group comprised 46% of the cohort but 76% of the total LYL. This pattern was consistent for all organs (Figure 9) In particular, lung recipients aged 50 years and older had the highest proportion of LYL to cancer (5.9% of expected life-years). Of the total LYL among lung recipients age 50+, 31% were related to lung cancer and 11% related to NHL. Heart recipients aged 50+ lost 4.2% of expected life-years to cancer, with a similar proportion of these LYL related to lung cancer (30%) and NHL (14%) as seen in lung recipients. Recipients under age 35 had the smallest proportion of total expected LYL across organs, ranging from 0.4% of expected life-years among kidney recipients to 1.2% among lung recipients. Of the LYL to cancer in this age group, 46% were associated with NHL.

		Method	Transplant recipients, n	Cancer cases, n (% of cohort)	Total LYL	LYL per transplant recipient: Mean	LYL per transplant recipient: 95% CI	Proportion of total life-years lost to cancer, %: Mean	Proportion of total life-years lost to cancer, %: 95% CI
Overall		Cox	221,962	13,074 (5.9%)	35,371	0.16	0.15, 0.17	1.9	1.8, 1.9
		KM			41,241	0.19	0.18, 0.19	2.2	2.0, 2.2
Transplanted organ									
	Kidney	Cox	136,534	7127 (5.2%)	19,148	0.14	0.13, 0.15	1.6	1.5, 1.6
		KM			22,200	0.16	0.16, 0.17	1.8	1.8, 1.9
	Liver	Cox	39,155	2366 (6.0%)	6890	0.18	0.16, 0.19	2.1	1.9, 2.2
		KM			7345	0.19	0.18, 0.20	2.2	2.1, 2.4
	Heart	Cox	22,912	2182 (9.5%)	5171	0.23	0.21, 0.24	2.8	2.5, 3.0
		KM			5659	0.25	0.23, 0.26	3	2.8, 3.3
	Lung	Cox	10,467	875 (8.4%)	2878	0.27	0.24, 0.31	4.4	3.9, 4.9
		KM			3012	0.29	0.26, 0.32	4.7	4.2, 5.1
	Other/ multiple	Cox	12,894	524 (4.1%)	1445	0.11	0.09, 0.13	1.3	1.1, 1.5
		KM			1519	0.12	0.09, 0.13	1.4	1.2, 1.6

LYL=life-years lost; CI=confidence interval; Cox=Cox regression model; KM=Kaplan-Meier survival estimates

Cox models used time since transplant as the time scale and were adjusted for sex, race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian/Pacific Islander, other), calendar year of transplant ((1987-1994, 1995-1999, 2000-2004, 2005-2008, 2009-2014), cancer registry, organ-specific covariates, and age at transplant. Since the effect of age on mortality was non-linear, age was modeled as a piecewise linear function with knots at age 5 and 19 years.

Table 19: Comparison of estimated life-years lost to cancer among solid organ transplant recipients using two approaches, one based on survival curves estimated from the Cox proportional hazards model and one based on Kaplan Meier (KM) survival curves

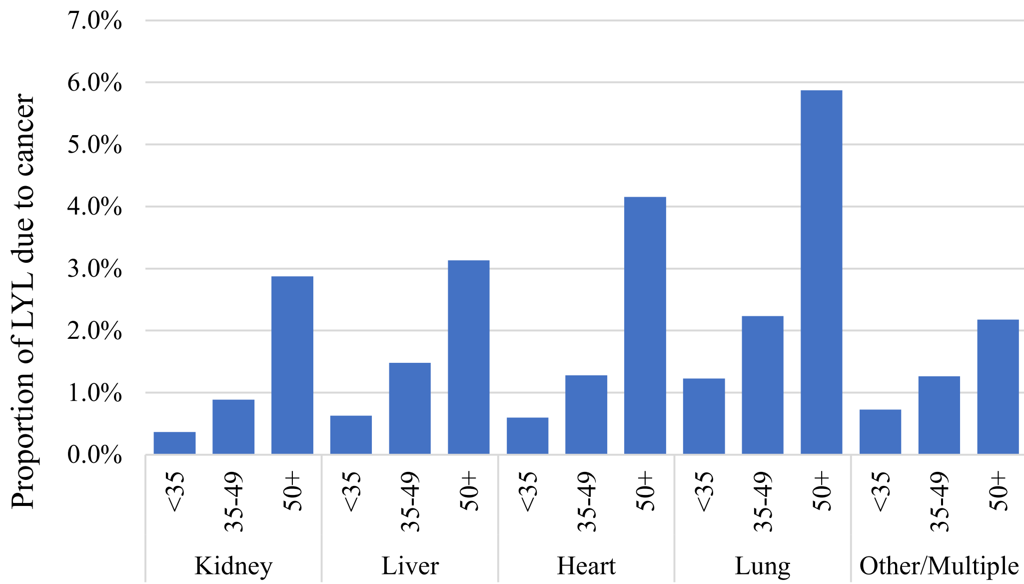


Figure 9: Proportion of expected life-years that were lost to cancer by age and transplanted organ

The mean LYL per person was higher for males compared to females (0.18 vs. 0.12 years, Table 17). This corresponded to 2.1% of total expected LYL due to cancer among males compared to 1.4% among female recipients. Non-Hispanic Whites had the highest mean LYL per person and proportion of total LYL to cancer (2.2%) followed by non-Hispanic Black recipients (1.6%). The proportion of total LYL to cancer for the other race/ethnic groups was lower (1.2-1.4% of expected life-years).

E. Discussion

Among US solid organ transplant recipients, an average of 0.16 years of life per person was lost to cancer within 10 years after transplant. Although this may seem like a small amount, the LYL per person translated into 1.9% of all life-years among this population expected in the absence of cancer. The total LYL depends on the proportion of recipients who develop cancer and the impact of cancer on mortality. The two cancer sites with the highest impact were lung cancer and NHL, which are common and associated with high mortality in transplant recipients.[83] These two cancer sites were responsible for 39% of the total LYL due to cancer in the cohort. Furthermore, once a recipient developed either lung cancer or NHL, they lost an average of almost five life-years.

Lung and heart recipients had the highest LYL to cancer. Moreover, lung cancer and NHL contributed a large portion of the total LYL in these recipients, especially among recipients age 50 and older. Lung and heart recipients receive the most intensive immunosuppression, which plays an important role in cancer incidence and mortality after transplantation.[18,83] NHL is a type of post-transplant lymphoproliferative disorder (PTLD) that results from poor immune control of Epstein-Barr virus (EBV) infection, and NHL risk is greatly elevated in transplant recipients due to the overall immunosuppression burden.[1] Among lung recipients (many of whom receive only a single transplanted lung), lung cancers most often arise in the native lung,[87] related to underlying lung disease and other risk factors, such as smoking, that put them at high risk for lung cancer. Other cancers contributed less to the overall LYL estimates (Table 18).

Although the proportion of LYL to cancer among recipients age 35 or younger was small (0.5% of all life-years), NHL was a large contributor. Among transplant recipients in this age group, NHL typically results from primary EBV infection.[68-70] It is important to note that the LYL that we estimated for young recipients were restricted by the censoring of follow-up that we imposed at 10 years after transplant. Indeed, the absolute LYL to cancer would be much larger for recipients if they were followed over a longer interval, and this would most substantially affect the estimates for younger recipients who have the longest potential lifetimes after transplant. Males had a higher proportion of LYL to cancer compared to females, partly reflecting their higher cancer incidence. Non-Hispanic White recipients had the highest proportion of LYL due to cancer compared to other races, although most differences by race/ethnicity were modest.

Several methods have been proposed to estimate LYL attributable to disease by comparing the mean survival times in two groups.[43-45] In our analyses, the mean survival times under the actual and hypothetical cancer-free scenarios were not independent, as they were each estimated using the same cohort of patients. We therefore used a bootstrap method that accommodated correlations to estimate the variances of the LYL. We obtained similar results using two approaches, namely, one based on Cox models that adjusted for baseline recipient demographic, medical, and transplant-related characteristics, and another sensitivity analysis that utilized a non-parametric comparison of areas under the two Kaplan-Meier curves.

Strengths of our study include its large sample of transplant recipients representative of the overall US transplant population, including recipients of all organ types. We obtained incident cancer information from population-based cancer registries that systematically ascertain cancers in their catchment areas. A limitation is that cancer registries do not collect information on non-melanoma skin cancers, so we could not include them in our analyses. These are common and cause substantial morbidity among transplant recipients.[56,57] One recent study estimated that non-melanoma skin and lip cancer were responsible for 15% of all cancer deaths among transplant recipients in Australia and New Zealand.[88] However, the study did not separate lip and skin cancers as causes of cancer-related death, and both are more common in those countries than the US due to high ultraviolet radiation exposure. Finally, some cancers may have been missed by the cancer registries in our study, if the linkage to the transplant registry did not match the cancer to the recipient, or if recipients migrated from the catchment area. However, outmigration was likely too uncommon to substantially affect our LYL estimates during a 10-year period after transplantation.[1]

The results of our study show that transplant recipients have a shortened life span after developing cancer compared to recipients who do not develop cancer. There are opportunities to reduce cancer mortality and extend the lives of transplant recipients through prevention and screening. Given that a large proportion of LYL are due to NHL and lung cancer, these cancers should be a priority. High levels of circulating EBV load can be used as a marker of PTLD risk, providing potential opportunities to screen for NHL, especially in pediatric recipients.[74] Following the recommendations for lung cancer screening in the general population, transplant providers should consider screening older transplant recipients with a smoking history using low-dose computed tomography.[89] Finally, cancer prevention strategies aimed at modifying lifestyle risk factors, especially smoking cessation,[90] should be prioritized to reduce cancer mortality.

In conclusion, our study demonstrates that cancer is associated with shortened lifespan among solid organ transplant recipients. Approximately 2% of all expected life-years in the first 10 years following transplantation are estimated to be lost to cancer. A large fraction of these LYL are associated with lung cancer and NHL. Further efforts aimed at cancer prevention, screening, and treatment for this population will be required.

7. Life-years lost due to cancer estimated using matching and comparison of results (Aim 2.1 and 2.3)

A. Introduction

In aim 2.1, the objective was to estimate LYL to cancer using matching. That is, recipients with cancer were matched to those who were cancer-free at the time of selection to serve as a cancer-free control group. Matching was conducted on the prognostic score with a sensitivity analysis using the propensity score.

B. Methods

The cohort was selected using data from the TCM study as described in chapter 0. In brief, first invasive cancers were identified from the linked cancer registries. Recipients with a prior cancer were excluded. Prior cancers were those diagnosed before transplant or within 90 days post-transplant. Follow-up thus started 90 days post-transplant, and ended at the earliest of death, loss to follow-up by the SRTR, end of cancer registry coverage, or 10 years post-transplant. Missing covariate values were imputed using multiple imputation.

The LYL due to cancer is estimated by the difference in mean survival between two groups: The group of recipients who developed cancer post-transplant and the group of controls who were similar to the cancer cases but represent a counterfactual cancer-free survival experience. The control group was selected by matching cases to controls using the prognostic score.

Prognostic scores are used to balance the baseline risk of the outcome between cancer and control recipients by summarizing the association between covariates and death.[39-41] That is, matching on the prognostic score creates a set of cases and controls with the same baseline risk of death. The prognostic scores were estimated from an organ-specific Cox model with covariates shown in Table 12 and Table 13. In these models, events were death from any cause. For recipients who developed cancer, the follow-up time was censored at the time of cancer diagnosis. For those who did not develop cancer, the follow-up time ended at loss to follow-up or end of cancer registry coverage. The prognostic score was estimated from each of the five imputed data sets and averaged to produce one score per recipient.

The matching criteria used to form the control group for each case were as follows: 1) controls must have the same transplanted organ and transplant year as the matched case; 2) controls must be within five years of age at transplant as the case; and 3) controls must be alive and cancer-free at the time of the case's cancer diagnosis. Controls can be matched to more than one case and may become a case at a later time. Up to five controls were selected that were within 0.2 difference of the prognostic score (i.e. caliper width) of the case. The number of controls available for different caliper widths is shown in Appendix 2.C.

The mean LYL per cancer was calculated as the difference in the mean survival of the cancer cases and matched controls. The mean survival for each group was estimated using a Kaplan-Meier curve with the follow-up time censored at 10 years post-transplant. The time scale origin was the date of cancer diagnosis for cancer cases or, for controls, the date of the matched case's cancer diagnosis. Separate Kaplan-Meier estimates were obtained for strata of case and control groups defined by the time since transplant to cases' cancer diagnosis in one year intervals. One-year groups from time since transplant were used for stratification (i.e., 0-1 years post-transplant, 1-2 years, …, 9-10 years).

This stratification was used since it accounted for the survival difference between those diagnosed shortly after transplant (blue line in Figure 10) compared to those diagnosed further from transplant (green line). The survival in the first few years following cancer diagnosis is very similar regardless of when the cancer occurred post-transplant. However, the survival for those diagnosed shortly after transplant is higher with time from cancer diagnosis. Although it is less pronounced, the same survival pattern is seen among the matched controls (Figure 11). Since the follow-up time was censored at 10 years post-transplant, recipients who were diagnosed 8 to 10 years post-transplant can only contribute a maximum of two years of follow-up. This is shown graphically by the green lines in Figure 10 and Figure 11. Consequently, only recipients who were diagnosed soon after transplant contribute to the estimation of end of the survival curve. Graphically this is shown by only the recipients diagnosed 0 to 4 years post-transplant contributing to the end of the survival curves. Since it does appear that there is a survival difference based on time of cancer diagnosis post-transplant, only using survival from those diagnosed soon after transplant to estimate the end of the survival curve may be biased. The stratification ensures that the survival comparison between cases and controls is controlled for time since transplant.

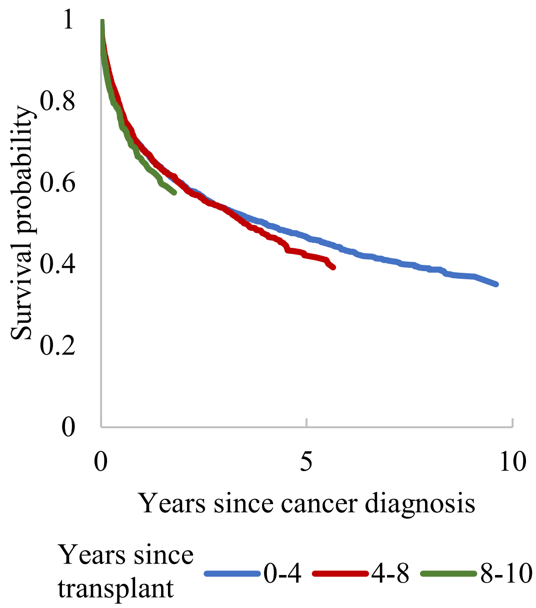


Figure 10: Kaplan-Meier curves for cancer cases stratified by time since transplant

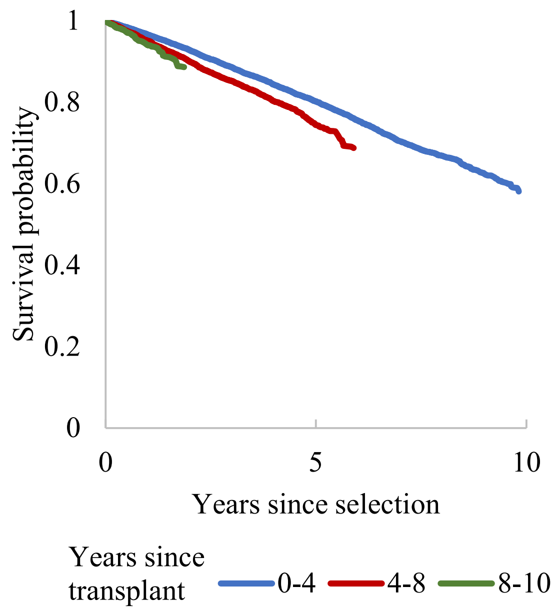


Figure 11: Kaplan-Meier curves for matched controls stratified by time since selection

A sample calculation on how the total LYL were estimated for kidney recipients in shown in Table 20. The first row of the table shows that there were 1260 cancer cases diagnosed 0 to 1 year after transplant and they were matched to 6265 controls. Since cancer diagnosis occurred within 1 year of transplant, they had a maximum follow-up time of 10 years. The mean survival time was estimated as the area under the Kaplan-Meier curves for cases and controls. The LYL per cancer is the difference in the mean survival and the total LYL in that stratum is the product of the LYL per cancer and the number of cancer cases. This was repeated for each stratum. Then the total LYL was obtained by summing the LYL over all strata. E.g. for kidney recipients, the total LYL was 13,875. Finally, the 95% confidence interval bounds were the 2.5th and 97.5th percentiles from the bootstrap distribution with 1000 bootstrap replicates.

Years since transplant to cancer diagnosis	Controls (n)	Cancer cases with matches (n)	Maximum follow-up time (years)	Mean survival		LYL per cancer	Total LYL
				Controls	Cases		
0 to <1	6265	1260	10	8.2	5.4	2.8	3534
1 to <2	5127	1031	9	7.4	4.6	2.7	2809
2 to <3	4439	894	8	6.6	4.1	2.6	2284
3 to <4	3939	795	7	5.9	3.6	2.2	1777
4 to <5	3698	745	6	5.1	3.3	1.8	1358
5 to <6	3221	651	5	4.4	2.9	1.5	969
6 to <7	2522	510	4	3.6	2.6	1.0	504
7 to <8	2329	473	3	2.8	2.0	0.7	347
8 to <9	1998	409	2	1.9	1.3	0.6	243
9 to 10	1569	321	1	1.0	0.8	0.2	51
Total	35,107	7089		4.7	3.1	2.0	13,875

LYL=Life-years lost

Table 20: Example LYL calculation using matching with time stratified by time since diagnosis/case selection for kidney recipients

Sensitivity analyses

In addition to the LYL estimated from the stratified Kaplan-Meier curves, several other methods were used after the matching. The first alternative was to use Cox models with delayed entry to the cohort instead of the stratified Kaplan-Meier curves. In this approach, separate Cox models were fit for cancer cases and for matched controls. The follow-up time was calculated as time from transplant until cohort exit or 10 years post-transplant, whichever came first. Cancer cases entered the cohort at the time of cancer diagnosis and controls entered at the time of selection. The mean life-years for cancer cases and controls was estimated by computing the area under the estimated baseline hazards.[85] In a second Cox model-based method, separate Cox models were again fit for the cancer cases and matched controls. However, here the time scale for the Cox models was calculated as the time from cancer diagnosis/selection to the end of follow-up or 10 years post-transplant without delayed entry or other adjustment.

Finally, the LYL per cancer case and total LYL were estimated from Kaplan-Meier curves using the propensity score as an alternative to the prognostic score for selecting controls. The propensity score was estimated from a Cox models with any cancer as the event. Death and loss to follow-up were censoring events. The time scale was the time since transplant. As with the prognostic score, the propensity score was estimated for each recipient and averaged from the five imputed data sets and used for matching.

C. Results

The estimates from the pooled Cox regression models used to estimate the prognostic score are shown in Table 21 through Table 23. A prognostic score was estimated for each organ recipient. The distributions of the prognostic scores are shown in Figure 12 for those who developed cancer and those who did not. The prognostic scores have good overlap for cases and their potential controls indicating that finding close matches to each case should be achievable.

	Kidney n=136534			Liver n=39155		
	HR	**95% CI**	**p-value**	**HR**	**95% CI**	**p-value**
Gender (ref: Male)						
Female	0.93	(0.90, 0.95)	<0.001	1	(0.95, 1.05)	<0.001
Race (ref: NH white)						
NH black	0.95	(0.91, 0.98)	<0.001	1.46	(1.36, 1.58)	
Hispanic	0.66	(0.63, 0.69)		0.91	(0.85, 0.97)	
API	0.59	(0.55, 0.63)		0.67	(0.59, 0.77)	
Other	0.9	(0.75, 1.07)		0.95	(0.72, 1.25)	

Table 21: Pooled estimates from the Cox model using imputed data for kidney and liver transplant recipients

Table 21 continued

	Kidney n=136534			Liver n=39155		
	HR	**95% CI**	**p-value**	**HR**	**95% CI**	**p-value**
Transplant year (ref: 1987-1994)						
1995-1999	1.18	(0.03, 51.9)	0.11	4.55	(2.85, 7.28)	0.38
2000-2004	0.93	(0.02, 42.8)		4.24	(2.65, 6.79)	
2005-2008	0.67	(0.01, 31.2)		4.6	(2.87, 7.38)	
2009-2014	0.49	(0.01, 22.6)		3.79	(2.34, 6.15)	
BMI (ref=underweight)						
Normal	0.8	(0.73, 0.88)	<0.001	0.76	(0.67, 0.86)	<0.001
Overweight	0.78	(0.71, 0.86)		0.7	(0.61, 0.80)	
Obese	0.89	(0.81, 0.97)		0.73	(0.64, 0.83)	
Diabetes (ref=No)						
Pre-1995	1.53	(0.03, 70.6)	<0.001	5.39	(3.38, 8.60)	<0.001
Yes	1.18	(1.10, 1.27)		1.51	(1.42, 1.60)	
Donor type (ref=Living)	1.35	(1.30, 1.40)	<0.001			
Time from initial dialysis to transplant (ref=None)						
0.01-0.99	1.51	(1.40, 1.63)	<0.001			
1-1.99	1.84	(1.70, 1.98)				
2-2.99	2	(1.85, 2.17)				
3-3.99	2.21	(2.03, 2.40)				
4-3.99	2.39	(2.18, 2.61)				
5+	2.9	(2.66, 3.16)				
Reasons for transplant (ref=No)						
Glomerular diseases	0.76	(0.72, 0.79)	<0.001			
Diabetes	1.83	(1.71, 1.95)	<0.001			
Polycystic kidneys	0.63	(0.59, 0.67)	<0.001			
Hypertensive nephrosclerosis	1.06	(1.01, 1.11)	0.02			
Other	0.89	(0.84, 0.94)	<0.001			
Viral hepatitis				1.35	(1.27, 1.44)	
Alcohol				1.17	(1.10, 1.25)	
PBC/PSC				0.65	(0.59, 0.71)	
Other				0.9	(0.84, 0.96)	

Baseline hazard was stratified in 1 year age groups.
The pooled type 3 p-value was combined using the method in Schafer, J.L. (1997) Analysis of Incomplete Multivariate Data. London: Chapman and Hall.

	Heart n=22912			Lung n=10467		
	HR	**95% CI**	**p-value**	**HR**	**95% CI**	**p-value**
Gender (ref: Male)	1.07	(1.01, 1.14)	0.02	1.01	(0.95, 1.08)	0.7
Female						
Race (ref: NH white)	1.78	(1.66, 1.90)	<0.001	1.06	(0.93, 1.20)	0.35
NH black	1.12	(1.02, 1.23)		1.12	(0.97, 1.29)	
Hispanic	1.08	(0.92, 1.26)		0.93	(0.69, 1.25)	
API	1.17	(0.83, 1.67)		1.27	(0.82, 1.96)	
Other						
Transplant year (ref: 1987-1994)	0.45	(0.06, 3.67)	0.04	1.09	(0.50, 2.35)	<0.001
1995-1999	0.41	(0.05, 3.44)		0.8	(0.36, 1.78)	
2000-2004	0.35	(0.04, 2.91)		0.79	(0.35, 1.77)	
2005-2008	0.29	(0.04, 2.42)		0.71	(0.32, 1.60)	
2009-2014						
BMI (ref=underweight)	0.94	(0.83, 1.07)	<0.001	0.92	(0.83, 1.02)	0.06
Normal	1.06	(0.93, 1.20)		0.99	(0.88, 1.11)	
Overweight	1.15	(1.01, 1.32)		1.03	(0.89, 1.18)	
Obese						
Diabetes (ref=No)	0.55	(0.07, 4.54)	<0.001	1.07	(0.48, 2.40)	
Pre-1995	1.37	(1.27, 1.47)		1.15	(1.03, 1.28)	
Yes	1.07	(1.01, 1.14)	0.02	1.01	(0.95, 1.08)	0.7
Tobacco use (ref=Yes)						
No				1.03	(0.92, 1.15)	0.57
Lung transplant (ref=double)				1.22	(1.13, 1.31)	<.0001
Single						
Reasons for transplant (ref=No)						
Cardiomyopathy	0.93	(0.85, 1.01)	0.08			
Coronary artery disease	1.14	(1.05, 1.25)	0			
Other	0.96	(0.85, 1.07)	0.45			
COPD				1.13	(0.99, 1.29)	0.06
Cystic fibrosis				0.96	(0.81, 1.14)	0.64
Idiopathic pulmonary fibrosis				1.06	(0.93, 1.20)	0.39
Other				1.03	(0.92, 1.15)	0.62

Baseline hazard was stratified in 1 year age groups.
The pooled type 3 p-value was combined using the method in Schafer, J.L. (1997) Analysis of Incomplete Multivariate Data. London: Chapman and Hall.

Table 22: Pooled estimates from the Cox model using imputed data for heart and lung transplant recipients

	Other/Multiple n=12894		
	HR	**95% CI**	**p-value**
Gender (ref: Male)	1.05	(0.97, 1.15)	0.23
Female			
Race (ref: NH white)	1.14	(0.99, 1.31)	0.22
NH black	0.92	(0.80, 1.07)	
Hispanic	1.05	(0.78, 1.41)	
API	0.91	(0.63, 1.33)	
Other			
Transplant year (ref: 1987-1994)	0.69	(0.22, 2.14)	<0.001
1995-1999	0.56	(0.18, 1.75)	
2000-2004	0.48	(0.15, 1.51)	
2005-2008	0.45	(0.14, 1.40)	
2009-2014			
BMI (ref=underweight)	0.7	(0.59, 0.84)	<0.001
Normal	0.67	(0.56, 0.81)	
Overweight	0.75	(0.59, 0.95)	
Obese			
Diabetes (ref=No)	0.54	(0.17, 1.71)	<0.001
Pre-1995	0.59	(0.52, 0.67)	
Yes	1.05	(0.97, 1.15)	0.23

Baseline hazard was stratified in 1 year age groups.
The pooled type 3 p-value was combined using the method in Schafer, J.L. (1997) Analysis of Incomplete Multivariate Data. London: Chapman and Hall.

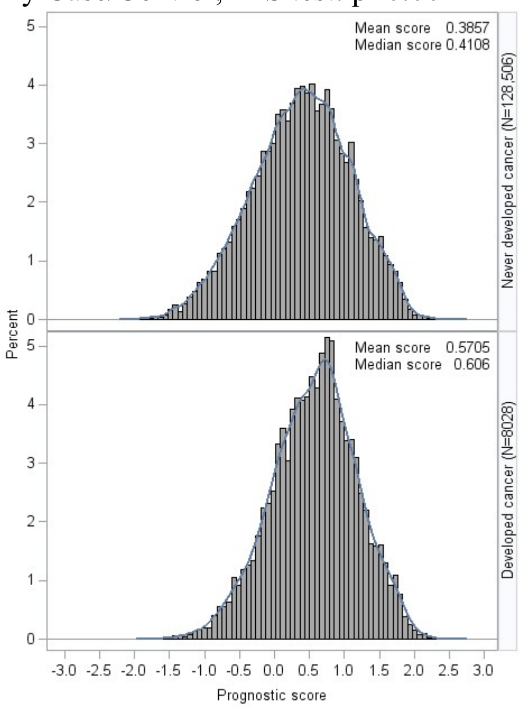
Table 23: Pooled estimates from the Cox model using imputed data for transplant recipients of other organ types

Table 24 shows the number of cancer cases, the number eligible for matching and the number of controls matched per case. For all organ types, very few cancer cases had no matched controls (≤3%), and most cases had the maximum of five matched controls.

The distributions of demographic characteristics for cancer cases and their matched controls are shown in Table 25 and Table 27. Overall, the cases and controls appeared to be well-matched based on the similarity of covariate distributions.

The LYL per cancer case and total LYL to cancer are shown in Table 28. The LYL per cancer case ranges from 2.0 years for kidney and lung recipients to 2.4 years among for liver recipients. The total LYL is highest among kidney recipients in part because this metric is dependent upon the number of cancer cases occurring in that subgroup. This measure will primarily be used to compare results from the various methods (Section D).

Kidney recipients
By Case/Control; K-S test: p<0.001



Liver recipients
By Case/Control; K-S test: p=0.03

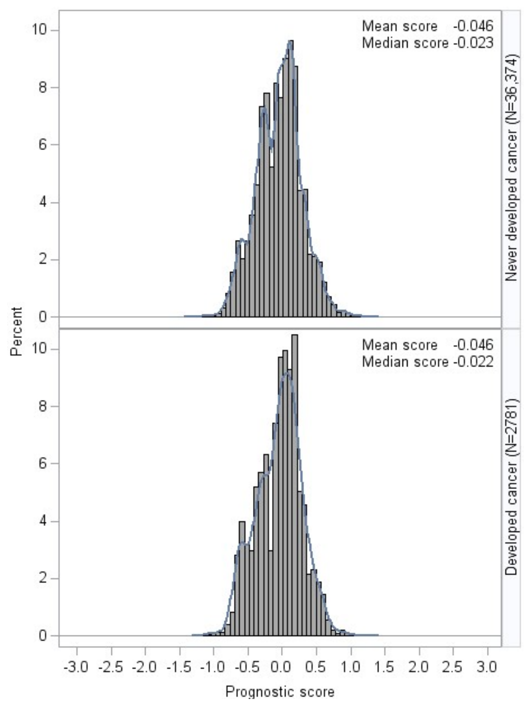
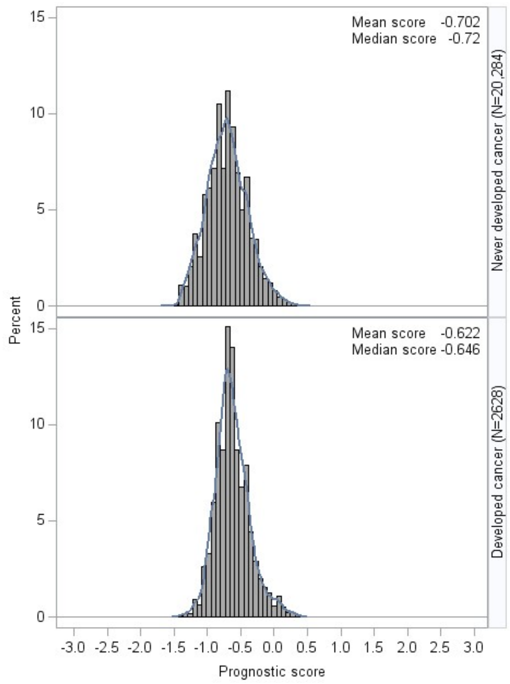


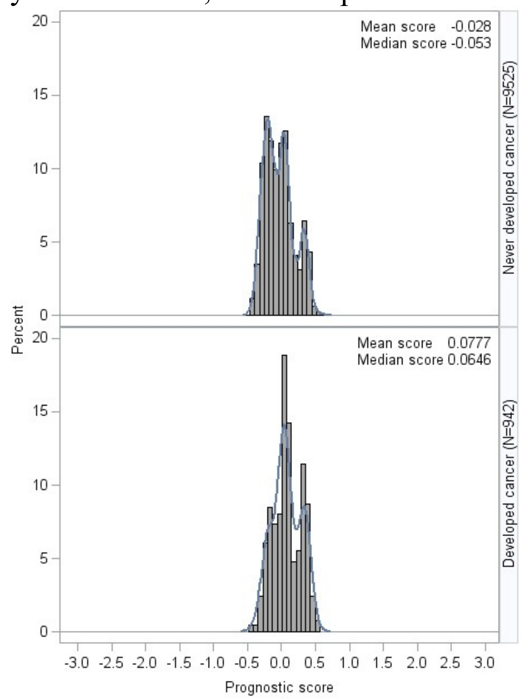
Figure 12: Estimates of prognostic score by cancer status in whole population before matching

Figure 12 continued

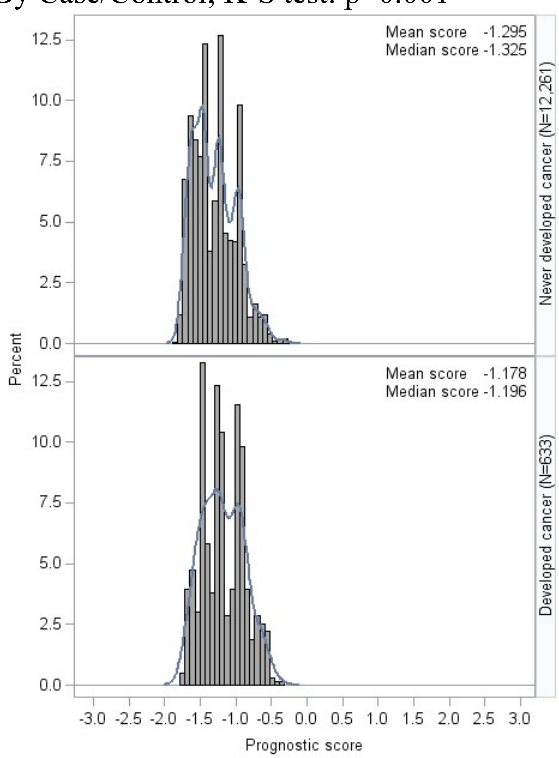
Heart recipients
By Case/Control; K-S test: p<0.001



Lung recipients
By Case/Control; K-S test: p<0.001



Other/multi recipients
By Case/Control; K-S test: p<0.001



							Controls per case								
	Total in cohort	Number of cancer cases	Number of cancers >10 years after transplant*	Number of cancers after excluding those >10 years	Total in analytic cohort	Number of cases	1	2	3	4	5	No controls		Total controls	Cases with controls
Kidney	136,534	8028	901	7127	136,370		28	43	32	33	6953	38	0%	35,107	7089
Liver	39,155	2781	415	2366	39,078		25	20	23	24	2257	17	1%	11,515	2349
Heart	22,912	2628	446	2182	22,466		20	27	31	26	2056	22	1%	10,551	2160
Lung	10,467	942	67	875	10,400		17	17	13	13	805	10	1%	4,167	865
Other/multi	12,894	633	109	524	12,874		21	18	21	18	428	18	3%	2,332	506

*These do not meet case definition and so they are used as potential controls

Table 24: Number of cancer cases and available matches within caliper width of 0.2

	Kidney				Liver			
	Cases		Controls		Cases		Controls	
Number of recipients:	7089		35107		2349		11515	
Phi								
Mean (SD)	0.56	0.61	0.56	0.60	-0.04	0.33	-0.04	0.32
Sex								
Male	4636	65%	21189	60%	1481	63%	7045	61%
Female	2453	35%	13918	40%	868	37%	4470	39%
Race								
NH white	4342	61%	20423	58%	1901	81%	9340	81%
NH black	1568	22%	7112	20%	142	6%	548	5%
Hispanic	810	11%	5059	14%	236	10%	1267	11%
API	328	5%	2325	7%	61	3%	317	3%
Other	41	1%	188	1%	9	0%	43	0%
Transplant Year								
1987-1994	1236	17%	6037	17%	407	17%	1969	17%
1995-1999	1986	28%	9871	28%	709	30%	3508	30%
2000-2004	2271	32%	11291	32%	794	34%	3891	34%
2005-2008	1295	18%	6424	18%	350	15%	1718	15%
2009-2014	301	4%	1484	4%	89	4%	429	4%
Age								
0-17	154	2%	735	2%	102	4%	493	4%
18-34	527	7%	2598	7%	75	3%	341	3%
35-49	1723	24%	8631	25%	579	25%	2893	25%
50-64	3409	48%	17071	49%	1295	55%	6448	56%
65+	1276	18%	6072	17%	298	13%	1340	12%
BMI								
Underweight	126	2%	679	2%	64	3%	228	2%
Normal	2221	31%	10834	31%	791	34%	3786	33%
Overweight	1958	28%	9667	28%	728	31%	3688	32%
Obese	1309	18%	6713	19%	538	23%	2858	25%
Unknown	1475	21%	7214	21%	228	10%	955	8%
Diabetes								
Yes	1578	22%	8677	25%	323	14%	1561	14%
No	3971	56%	18902	54%	1538	65%	7709	67%
Unknown	304	4%	1491	4%	81	3%	276	2%
Unknown (before 1995)	1236	17%	6037	17%	407	17%	1969	17%

Table 25: Comparison of characteristics of cases and controls for kidney and liver recipients based on top 5 matches and caliper width of 0.2

Table 25 continued

	Kidney				Liver			
	Cases		Controls		Cases		Controls	
Donor type								
Deceased	4807	68%	24148	69%				
Living	2282	32%	10959	31%				
Dialysis								
No dialysis	676	10%	3035	9%				
0-1	1464	21%	7519	21%				
1-2	1553	22%	7552	22%				
2-3	1007	14%	5252	15%				
3-4	701	10%	3607	10%				
4-5	458	6%	2477	7%				
5+	1164	16%	5328	15%				
Unknown	66	1%	337	1%				
Reasons for transplant								
Glomerular diseases	1903	27%	8971	26%				
Diabetes	1585	22%	8912	25%				
Polycystic kidneys	785	11%	4015	11%				
Hypertensive nephrosclerosis	1587	22%	7509	21%				
Other	962	14%	4532	13%				
Viral hepatitis					703	30%	3780	33%
Alcohol use					822	35%	3725	32%
PBC/PSC					405	17%	1766	15%
Other					730	31%	3606	31%

SD=standard deviation, BMI=body mass index, PBC=primary biliary cirrhosis, PSC=primary sclerosing cholangitis

	Heart				Lung			
	Cases		Controls		Cases		Controls	
Number of recipients:	2160		10051		865		4167	
Phi								
Mean (SD)	-0.63	0.25	-0.64	0.24	0.07	0.21	0.07	0.21
Sex								
Male	1795	83%	8791	83%	520	60%	2217	53%
Female	365	17%	1760	17%	345	40%	1950	47%
Race								
NH white	1785	83%	9012	85%	773	89%	3718	89%
NH black	207	10%	782	7%	48	6%	249	6%
Hispanic	122	6%	508	5%	26	3%	151	4%
API	39	2%	212	2%	13	2%	33	1%
Other	7	0%	37	0%	5	1%	16	0%
Transplant Year								
1987-1994	568	26%	3732	35%	98	11%	448	11%
1995-1999	747	35%	3672	35%	213	25%	1024	25%
2000-2004	552	26%	2728	26%	302	35%	1483	36%
2005-2008	246	11%	1192	11%	194	22%	941	23%
2009-2014	47	2%	227	2%	58	7%	271	7%
Age								
0-17	89	4%	428	4%	13	2%	52	1%
18-34	45	2%	220	2%	52	6%	241	6%
35-49	260	12%	1223	12%	120	14%	542	13%
50-64	1461	68%	7294	69%	549	63%	2795	67%
65+	305	14%	1386	13%	131	15%	537	13%
BMI								
Underweight	74	3%	317	3%	92	11%	383	9%
Normal	853	39%	4351	41%	357	41%	1908	46%
Overweight	740	34%	3566	34%	284	33%	1316	32%
Obese	283	13%	1396	13%	100	12%	419	10%
Unknown	210	10%	921	9%	32	4%	141	3%
Diabetes								
Yes	300	14%	1414	13%	71	8%	282	7%
No	1224	57%	6229	59%	678	78%	3335	80%
Unknown	68	3%	176	2%	18	2%	102	2%
Unknown (before 1995)	568	26%	2732	26%	98	11%	448	11%

Table 26: Comparison of characteristics of cases and controls for heart and lung based on top 5 matches and caliper width of 0.2

Table 26 continued

	Heart				Lung			
	Cases		Controls		Cases		Controls	
Tobacco use								
Yes	533	62%	2542	61%				
No	142	16%	746	18%				
Lung transplant								
Single	553	64%	2628	63%				
Double	312	36%	1539	37%				
Reasons for transplant								
Cardiomyopathy	952	44%	4403	42%				
Coronary artery disease	1258	58%	6441	61%				
Other	124	6%	440	4%				
COPD					49%	2236	54%	49%
Cystic fibrosis					8%	288	7%	8%
Idiopathic pulmonary fibrosis					28%	1018	24%	28%
Other					19%	778	19%	19%

	Other/multiple			
	Cases		Controls	
Number of recipients:	506		2332	
Phi				
Mean (SD)	-1.22	0.29	-1.24	0.27
Sex				
Male	305	60%	1385	59%
Female	201	40%	947	41%
Race				
NH white	382	75%	1937	83%
NH black	55	11%	181	8%
Hispanic	57	11%	157	7%
API	10	2%	42	2%
Other	2	0%	15	1%
Transplant Year				
1987-1994	81	16%	367	16%
1995-1999	143	28%	664	28%
2000-2004	168	33%	786	34%
2005-2008	87	17%	406	17%
2009-2014	27	5%	109	5%
Age				
0-17	24	5%	81	3%
18-34	85	17%	396	17%
35-49	258	51%	1230	53%
50-64	130	26%	601	26%
65+	9	2%	24	1%
BMI				
Underweight	23	5%	78	3%
Normal	237	47%	1262	54%
Overweight	119	24%	588	25%
Obese	44	9%	174	7%
Unknown	83	16%	230	10%
Diabetes				
Yes	323	64%	1578	68%
No	95	19%	366	16%
Unknown	7	1%	21	1%
Unknown (before 1995)	81	16%	367	16%

Table 27: Comparison of characteristics of cases and controls for recipients of other organs based on top 5 matches and caliper width of 0.2

Organ transplanted	Controls (n)	Cancer cases (n)	LYL per case	Total LYL among cohort	95% CI	
Kidney	35,107	7,089	2.0	13,875	13,218	14,565
Liver	11,515	2,349	2.4	5,639	5,230	6,063
Heart	10,551	2,160	2.2	4,380	3,984	4,692
Lung	4,167	865	2.0	1,905	1,651	2,120
Other	2,332	506	2.3	1,141	915	1,318

LYL: Life-years lost, CI=confidence interval

Table 28: Life-years lost to cancer by organ using the matching method with stratified Kaplan-Meier curves

Sensitivity analyses

The results from the stratified Kaplan-Meier method and the Cox regression models using the time scale since cases' cancer diagnosis/selection were the most similar, especially when accounting for the 95% confidence intervals (Table 29). Two potential contributors to the differences are the censoring and the risk sets constructed by each method. Since all methods incorporate censoring at 10 years post-transplant this is unlikely a source of any differences in the results. So the risk sets are the likely cause of these differences in LYL estimates. The risk set in the stratified Kaplan-Meier approach is comprised of the cancer cases diagnosed in each stratum (e.g. diagnosed 0 to 1 years after transplant) and their matched controls who survived to time *t,* where *t* is the time from the cases' transplant to cancer diagnosis. The risk set for the Cox models with time since transplant as the time scale (delayed entry) are cancer cases and controls who have survived to time *t* after transplant. Finally, the risk sets in the Cox models with time since cancer diagnosis as the time scale includes the cancer cases and their matched controls who survived to time *t,* but does not take into account when the cancer occurred post-transplant. So this method constructs risk sets similar to the stratified approach but compares cancers and controls who are a mix of when the cancer occurred.

Organ transplanted	Cancer cases with matches (n)	Controls (n)	Total LYL among cohort	95% CI	
Kidney					
Stratified KM	7089	35,107	13,875	13,218	14,565
Cox models w/time since transplant	7089	35,107	42,022	39,358	44,876
Cox models w/time since cancer dx/selection	7089	35,107	20,788	19,738	21,658
Liver					
Stratified KM	2349	11,515	5,639	5,230	6,063
Cox models w/time since transplant	2349	11,515	14,758	13,084	16,417
Cox models w/time since cancer dx/selection	2349	11,515	8,687	6,959	9,445
Heart					
Stratified KM	2160	10,551	4,380	3,984	4,692
Cox models w/time since transplant	2160	10,551	13,874	12,004	15,702
Cox models w/time since cancer dx/selection	2160	10,551	7,158	6,640	7,634
Lung					
Stratified KM	865	4167	1,905	1,651	2,120
Cox models w/time since transplant	865	4167	4,016	3,377	4,712
Cox models w/time since cancer dx/selection	865	4167	2,502	2,208	2,792
Other/Multi					
Stratified KM	506	2332	1,141	915	1,318
Cox models w/time since transplant	506	2332	3,109	2,601	3,698
Cox models w/time since cancer dx/selection	506	2332	1,767	1,499	2,035

CI=Confidence interval, KM=Kaplan-Meier, dx=diagnosis

Table 29: Comparison of life-years lost to cancer from each method after matching

Since the risk sets for the stratified Kaplan-Meier approach and the Cox models with time since cancer diagnosis/selection are similar except for the control over when the cancer developed post-transplant, it is not surprising that these results are the most similar. The different results using the Cox models with time since transplant as the time scale may perhaps be explained by an interaction happening between the two time scales and when deaths occur. I have chosen the results using the stratified Kaplan-Meier method as the primary results, because it provides some control as to when the cancer occurred post-transplant.

Finally, the LYL using the propensity score for matching compared to the prognostic score are shown in Table 30. In general, more cases were matched using the propensity score. However, there were no substantive differences in results between the two scores.

	Score used for matching	Number cancer cases with matches	LYL per case	Total LYL among cohort	95% CI	
Kidney						
	Prognostic	7089	2.0	13,875	13,218	14,565
	Propensity	7117	2.0	14,005	13,259	14,703
Liver						
	Prognostic	2,349	2.4	5,639	5,230	6,063
	Propensity	2,361	2.4	5,679	5,225	6,048
Heart						
	Prognostic	2,160	2.0	4,380	3,984	4,692
	Propensity	2,176	2.1	4,462	4,072	4,797
Lung						
	Prognostic	865	2.2	1,905	1,651	2,120
	Propensity	855	2.2	1,900	1,619	2,082
Other/Multi						
	Prognostic	506	2.3	1,141	915	1,318
	Propensity	519	2.3	1,216	976	1,370

Table 30: Comparison of life-years lost from matching controls on prognostic vs. propensity scores

D. Comparison and discussion of methods

Table 31 shows results from using the matching method and the modeling method to construct the cancer-free experience. Here the results from stratified Kaplan-Meier are shown since they were selected as the primary results from the matching method. For the modeling method, results from both the Kaplan-Meier and Cox models are shown.

The differences between the matching method results are discussed above. Among the modeling methods, the Cox models resulted in slightly lower estimates of total LYL compared to using the Kaplan-Meier curves. The main difference between the two methods is that the Cox models were adjusted for covariates whereas Kaplan-Meier curves are not adjusted for covariates.

The covariate adjustment in the Cox regression models addresses potential confounding of the risk of death. However, the quantity of interest here is the LYL. This quantity is the difference in mean survival of the cancer cases with their hypothetical cancer-free experience. If the cancer cases and the recipients who comprise the cancer-free experience differed on some covariate distribution, then the LYL estimate could be confounded. However, the same recipients and their baseline data are used to estimate the survival from the cancer and cancer-free models. Therefore, the survival curves from the cancer and cancer-free models cannot be confounded. In fact, since they are the same recipients at baseline the estimates are completely unconfounded because their covariate distributions are exactly the same.

Finally, comparing the results from the modeling method using Cox regression models to the results using the stratified Kaplan-Meier with matching shows that the total LYL to cancer is higher when estimated using the modeling method. These differences are discussed in the Discussion (Chapter 8).

Organ transplanted	Total LYL among cohort	95% CI	
Kidney			
Matching, Stratified KM	13,875	13,218	14,565
Modeling, KM	22,200	21,379	23,037
Modeling, Cox model	19,148	18,052	19,983
Liver			
Matching, Stratified KM	5,639	5,230	6,063
Modeling, KM	7,345	6,918	7,811
Modeling, Cox model	6,890	6,326	7,393
Heart			
Matching, KM	4,380	3,984	4,692
Modeling, KM	5,659	5,283	6,061
Modeling, Cox model	5,187	4,765	5,577
Lung			
Matching, Stratified KM	1,905	1,651	2,120
Modeling, KM	3,012	2,721	3,321
Modeling, Cox model	2,878	2,546	3,211
Other/Multi			
Matching, Stratified KM	1,141	915	1,318
Modeling, KM	1,519	1,286	1,736
Modeling, Cox model	1,453	1,189	1,701

LYL=life-years lost, CI=confidence interval, KM=Kaplan-Meier

Table 31: Comparison of the total LYL and proportion of LYL from the selected matching method and modeling methods

8. Discussion

The objective of this dissertation was to compare several population-based mortality metrics for the US transplant population. The estimates provided herein have not previously been reported for this population. Moreover, novel biostatistical methods were utilized and several approaches were taken to estimate each statistic.

Cancer is a large contributor to mortality among solid organ transplant recipients. I found that 13% of deaths were attributable to cancer, corresponding to a cancer-attributable mortality rate of 516 per 100,000 person-years. Furthermore, cancer results in an average of 2.7 years of life lost per recipient with cancer within the first 10 years post-transplant. These results vary by population subgroups and cancer sites. Based on cancer site specific results, I infer that immunosuppression likely plays a large role in cancer mortality among organ transplant recipients.

A. Risk factors for cancer

Several characteristics of organ transplant recipients and their post-transplant treatment place these patients at increased risk of cancer. Transplant recipients may have health behaviors known to increase cancer risk in the general population, such as cigarette smoking, obesity, lack of physical activity, and dietary factors such as high alcohol intake.[14] Indeed, smoking history has been associated with increased cancer risk among kidney and liver recipients.[91,92] While no studies have explicitly examined an association between lack of physical activity and cancer risk in the transplant population, recipients have been found to have lower levels of physical activity which may increase cancer risk as in the general population.[93-97] Also, alcohol consumption is associated with increased risk of several cancers in the general population and is associated with worse outcomes among liver recipients.[98,99]

Underlying chronic disease that contributed to organ failure may affect the native organ left in place among some kidney and lung recipients. The majority of kidney cancers arise in the native kidney of kidney recipients and may be due to kidney diseases such as acquired cystic kidney disease.[16,100-102] Similarly, the majority of lung cancers among lung recipients develop in the native lung,[17,103] and are likely caused by smoking-related lung disease such as chronic pulmonary obstructive disease,[17] or other lung conditions such as idiopathic pulmonary fibrosis.[104]

Immunosuppression is a hallmark of post-transplant treatment and contributes to cancer development. This is particularly true for cancers caused by viruses as Hodgkin an non-Hodgkin lymphoma (Epstein-Barr virus), liver cancer (hepatitis C and B viruses), ano-genital cancers (human papillomavirus), and Kaposi sarcoma (human herpesvirus 8).[22] The duration and intensity of immunotherapy influences the development of malignancy after transplant even for non-infection related cancers. Thus, heart and lung recipients who generally receive very intensive immunosuppression have a higher risk of cancer compared to kidney recipients.[1,18] Cancers that develop post-transplant may also be more aggressive exhibiting accelerated growth and spread.[18,105,106] For example, melanomas in recipients present at more advanced stages and are associated with worse survival compared to non-immunocompromised persons.[106]

B. Cancer incidence and mortality in the transplant population

Solid-organ transplant recipients have an increased risk for developing cancer compared with the general population.[1,66,67,107-112] A population-based study in the US from 1987 to

2008 found cancer incidence among transplant recipients was 1375 per 100,000 person-years, which corresponded to an absolute excess risk of 719 per 100,000.[1] This corresponded to about a two-fold higher risk of developing cancer compared to the general population.

Furthermore, cancer incidence varies by transplanted organ. A registry study conducted in the UK found that risk of cancer was about 2-fold higher for recipients of kidney and liver transplants and 3 times higher in recipients of heart or lung transplants.[67] A meta-analysis of 21 studies found similar results with a 3-fold increased risk of cancer among heart recipients and 4-fold increased risk among lung recipients.[107]

Solid-organ transplantation has been associated with increased risk of both infection and non-infection related cancers, although risk is substantially higher for infection-related cancers.[1,66,109,110] For example, in analyses using population-based data with recipients transplanted from 1987 to 2008, the incidence of non-Hodgkin lymphoma was elevated 7.5-fold compared to the general population, the risk of penile and vulva cancers was also elevated.[1] Similar results were found in a meta-analysis with incidence significantly increased for cancers with known or suspected infectious cause.[111] Also, even after excluding cancers within 6 months from transplant, the risk of liver cancer was elevated 1.5-fold among liver recipients.[23] Risks of several non-infection related cancers were elevated such as a two-fold increase in lung cancer, 4.6-fold increase in kidney cancer, 2.4-fold increase in melanoma, and a 3-fold increase in kidney cancer, among others.

In addition to cancer incidence, cancer mortality is an important measure of burden since it reflects the final downstream outcome in a population accounting for cancer incidence, treatment, and survival following a cancer diagnosis. As the population of transplant recipients grows, the number of recipients who develop cancer and subsequently die from cancer will also increase making this an important public health concern. Thus, a comprehensive understanding of the cancer mortality burden among recipients is critical.

Cancer mortality can be measured using different statistics, each providing a different view of the cancer mortality burden. One metric is the standardized mortality ratio which estimates the risk of cancer death among recipients compared to the general population. This statistic informs which cancers cause excess risk of death beyond what is expected among the general population. Indeed, several studies conducted outside of the US have shown increased cancer mortality among recipients of kidney,[2-4] liver, heart and lung transplants.[2,5] These studies have consistently shown that risk of cancer death is 2 to 3 times more likely among transplant recipients compared to the general population. Specifically, results from a recent Canadian study of kidney, liver, heart and lung recipients showed that the mortality risk from cancer was higher than the general population (SMR: 2.84, 95% CI: 2.18, 3.07).[2] In contrast, the only previous study in the US to estimate excess cancer mortality did not find an increased risk of cancer death among kidney recipients (SMR: 0.96, 95% CI: 0.92, 1.0).[24] However, this study relied on cause of death reported to the US Renal Data System which was unknown for 41% of the transplant recipients. These deaths were classified as non-cancer deaths which leaves a substantial opportunity for misclassification bias.

The cancer mortality burden can also be characterized by the cancer-specific mortality rates. These statistics are useful to compare subpopulations and trends over time. In a population-based study using SRTR data, the cancer-specific mortality rate was 114 per 1,000 person-years.[12] Mortality rates varied greatly by cancer site and were highest for pancreatic and lung cancers (1166 and 538 per 1,000 person-years, respectively). In contrast to relative measures such as the SMR, mortality rates estimate the number of persons who develop cancer over time.

In addition, other dimensions of the cancer mortality burden can be estimated. One such metric is the proportion of deaths attributed to cancer. This relative metric, estimated using the population-attributable fraction, summarizes how much of the mortality burden is due to cancer relative to other causes of death. This measure is useful for identifying particular cancers that may be contributing to cancer mortality and these may be the cancers to target for enhanced screening or prevention. Finally, mortality may also be summarized by the life-years lost to cancer. This is primarily used to evaluate the economic and social burden of cancer since it captures how much life was shortened by cancer. This dissertation is the first study to provide estimates of the fraction of deaths attributable to cancer and the life-years lost to cancer among recipients. This addresses a research gap and augments other published cancer mortality statistics.

C. Statistical methods

In my dissertation, I conducted two descriptive epidemiology studies utilizing novel biostatistical methods to quantify the burden of cancer mortality among US solid organ transplant recipients. Three dimensions of the cancer mortality burden were estimated: Proportion of deaths due to cancer, cancer-attributable mortality rate, and life-years lost due to cancer. Each quantity was estimated using two different statistical approaches.

The proportion of deaths due to cancer among transplant recipients, also known as the population attributable fraction, is a relative measure of impact since it is a function of the risk for non-cancer causes of death. This measures the extent to which cancer causes death in the whole population in relation to overall mortality. This measure gives an indication of the cancer's relative ranking compared to other individual or combined causes of death. The attributable fraction among those with cancer is represented in Figure 13. The all-cause mortality rate among recipients who never developed cancer is shown in the left bar and the mortality rate among receipts who developed cancer is shown in the right bar. The difference in the mortality rates (shaded in blue) is the excess mortality attributed to incident cancer. An assumption is that the two groups are similar except for cancer incidence (i.e. the association between cancer and mortality is unconfounded).

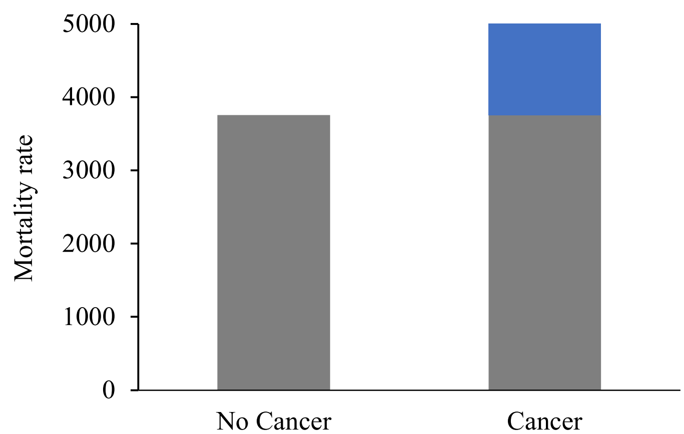


Figure 13: Graphical representation of the population attributable fraction among those with cancer

Mathematically, the attributable fraction is calculated as shown in , which makes it clear that the attributable fraction is a function of the relative risk of cancer. Notably, this attributable fraction will be close to 1 when the relative risk of death is large. In order to control for potential confounding in the estimation of the relative risk, the adjusted hazard ratio (HR) estimated from a Cox model was used.

$$\text{Attributable fraction}_{cancer}(\text{AF}_{cancer}) = \frac{\text{Mort rate}_{\text{Cancer}} - \text{Mort rate}_{\text{No cancer}}}{\text{Mort rate}_{\text{Cancer}}}$$

$$= (RR_d\text{-}1)/RR_d$$

where Mort rate=all-cause mortality rate per 100,000 person-years among recipients with cancer and without and RR_d=risk of death associated with a cancer diagnosis

Equation 3: Attributable fraction

The population attributable fraction is then obtained by multiplying the attributable fraction by the percentage of deaths with a prior cancer diagnosis (Equation 4).

$$\text{Population attributable fraction (PAF)} = \text{AF}_{\text{cancer}} * p_d$$

where p_d= percentage of deaths with a prior cancer diagnosis

Equation 4: Population attributable fraction

Finally, the cancer-attributable mortality rate can be estimated from the PAF and mortality rate among those from cancer (). The PAF determines how much of the cancer mortality rate is due to cancer.

$$\text{Cancer-attributable mortality rate} = \text{PAF} \times \text{Mort rate}_{\text{cancer}}$$

$$= \text{PAF} \times \frac{\text{Total number of deaths}}{\text{Person-years}_{\text{Transplant cohort}}}$$

Equation 5: Cancer-attributable mortality rate

The LYL estimates the extent to which life is shortened by cancer.[26] Figure 14 is a schematic of this quantity. The actual lifetime of a recipient is observed from the time of transplant through cancer diagnosis to subsequent death. I assume that this recipient's lifetime would have been longer had no cancer occurred. This is shown by the counterfactual lifetime that cannot be observed. The life-years lost to cancer is the difference between the actual lifetime and the counterfactual lifetime (shaded box).

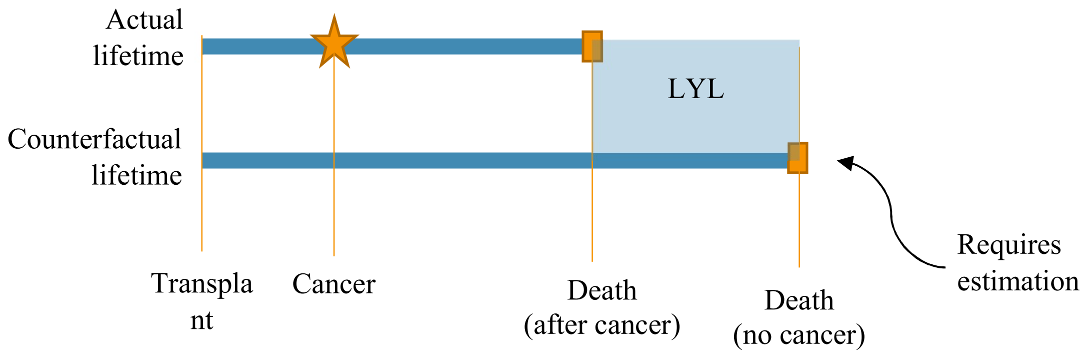


Figure 14: Schematic showing estimation of life-years lost

Statistically, the mean LYL due to cancer is estimated by comparing the survival curve of the overall cohort, which is an estimate of the actual lifetime, and the hypothetical cancer-free cohort, which represents the counterfactual lifetime. In this dissertation, the area between the two survival curves was estimated from Cox models and then using Kaplan-Meier curves in a sensitivity analysis. In contrast to mortality rates, which quantify the number of deaths occurring in a population, LYL estimates are often used to evaluate the social and economic burden that result from death.

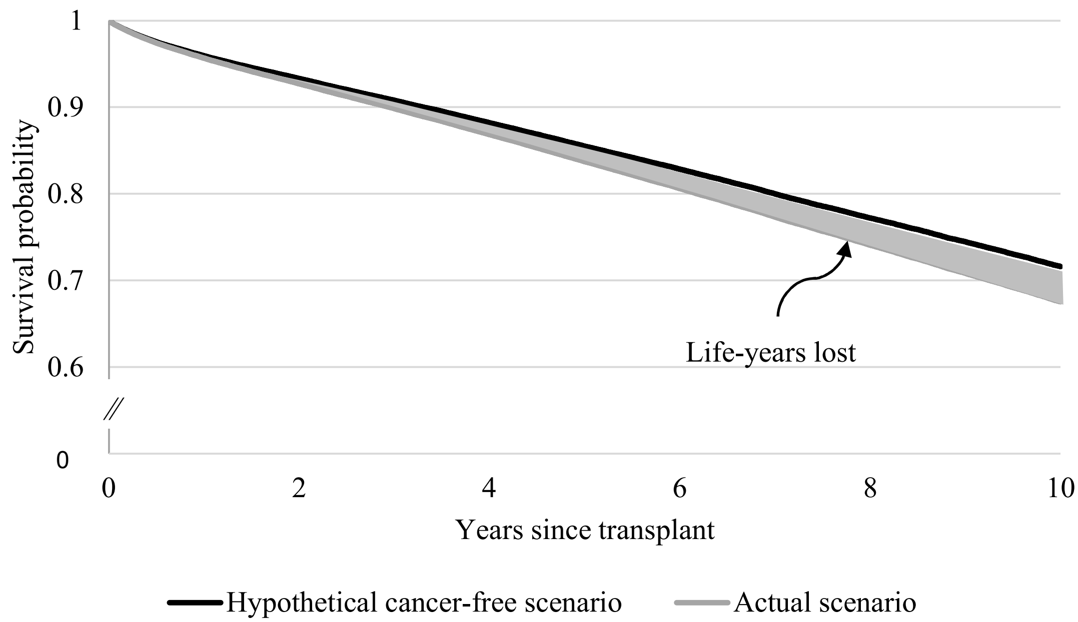


Figure 15: Life-years lost due to cancer among transplant recipients

The proportion of deaths due to cancer, the cancer-attributable mortality rate and the LYL to cancer were estimated using data from the National Cancer Institute's Transplant Cancer Match Study, which is a linkage of the US Scientific Registry of Transplant Recipients (SRTR) and 17 population-based state and regional cancer registries. SRTR data include recipient demographic and transplant characteristics, vital status and organ function. COD is based on clinical information reported by transplant centers and updated over time, so recipients may have multiple CODs. The cancer registries provided data on incident cancers (excluding non-melanoma skin cancer) and COD covering approximately half of the US transplant population. The study cohort included 221,962 solid organ transplant recipients and 13,074 incident cancers.

D. Population-attributable fraction and cancer-attributable mortality rate

My dissertation is the first study to estimate the proportion of deaths due to cancer and the cancer-attributable mortality rate among recipients. In my first study, I found that 13% of deaths were attributable to cancer, corresponding to a cancer-attributable mortality rate of 516 per 100,000 person-years. Lung cancer was the largest contributor to mortality (PAF=3.1%), followed by non-Hodgkin lymphoma (NHL, PAF=1.9%), colorectal cancer (PAF=0.7%), and kidney cancer (PAF=0.5%). The PAF depends on cancer incidence and its impact on mortality. The two cancer sites with the highest PAFs (lung cancer and NHL) are both common and associated with high mortality in transplant recipients. Although pancreatic cancer was highly fatal, it was not very common and so did not result in a high PAF (0.5%).

Some of the findings are likely related to differences in immunosuppression. For example, I found that lung and heart recipients, who generally receive intensive immunosuppression, had the highest cancer-attributable mortality (121 and 75 per 100,000 person-years, respectively). Lung and heart recipients had high PAFs for NHL, a cancer strongly associated with EBV (PAF=1.7% and 2.4%, respectively). These recipients also had high PAFs for lung cancer, which is linked to smoking and, among lung recipients, end-stage lung disease.[17] Among the overall cohort, the cancer-attributable mortality rate also increased with time since transplant, and lung cancer and NHL contributed prominently greater than five years post-transplant. This pattern partly reflects rising incidence of these cancers with longer duration since transplant and prolonged exposure to immunosuppressant medications.[1,66,67]

Cancer-attributable mortality rates increased with age at transplantation, reaching 1229 per 100,000 person-years among recipients 65+ years old. NHL was the largest contributor among children (PAF=4.1%). NHL, when it occurs in transplant recipients, is a type of post-transplant lymphoproliferative disorder (PTLD), which among young recipients results from primary EBV infection.[68-70] Lung cancer was a large contributor among 50+ year-olds (PAFs=3.7-4.3%). Heart recipients had the highest PAF (16.4%), but lung recipients had the highest cancer-attributable mortality rate (1241 per 100,000 person-years). Overall, mortality attributable to cancer increased steadily with longer time since transplant, reaching 15.7% of deaths (810 per 100,000 person-years) 10+ years post-transplant. This pattern partly reflects rising incidence of these cancers with longer duration since transplant and prolonged exposure to immunosuppressant medications.[1,66,67]

The interpretation of the PAF relies on a causal relationship between cancer and excess mortality, and a key assumption is that the association between cancer and death is estimated without confounding. The Cox models were adjusted for recipient demographic characteristics. However, if there was substantial residual confounding, then PAFs would not accurately reflect mortality that should be attributed to cancer. One approach to assessing the impact of a potential confounder is to note that the PAF is a function of (HR-1)/HR, where HR is the hazard ratio estimating the risk of death among those with cancer compared to those without cancer. Since the association between cancer and death is generally strong (i.e. a large HR), this component of the PAF formula approaches 1 and the PAF is very close to the proportion of deaths preceded by a cancer diagnosis (i.e. p_d). A potential confounder would need to have a profound reduction on the HR between cancer and death to substantially shift the PAF results.

I conducted a sensitivity analysis to assess the potential impact of tobacco use which is likely to be the largest unmeasured confounder for lung cancer mortality estimates. I estimated the bias factor to adjust the HR for smoking.[71] The HR further adjusted for tobacco use was 6.3 and, although this was lower than the HR of 10.4 found in the primary analysis, the resulting PAFs were similar. Specifically, the PAF for lung cancer was 2.5% compared to 3.1% found in the primary analysis. Therefore, we conclude that the PAF for lung cancer is not greatly affected by unmeasured confounding by tobacco use. While there may be other unmeasured confounders, it is unlikely that any are so large as to clinically change the PAF results.

Finally, PAF is a relative metric and useful to compare the impact of cancer on death relative to other causes. However, it is important to consider the PAF in the context of overall mortality. About 21% of deaths in the US general population were associated with cancer in 2017.[113] This is higher than the estimate of 13% among recipients. One reason for a lower PAF among recipients is that they have other competing risks of death either from the transplant itself or other health conditions and so the proportion due to cancer is lower. However, transplant recipients have a substantially higher cancer-attributable mortality than the US general population (516 vs. 153 per 100,000). This indicates that while cancer mortality is elevated among recipients, it may comprise a smaller piece of the total overall mortality because recipients have many other competing risks of death.

E. Cancer-attributable mortality using cause of death

An alternative approach is to use COD information to estimate the cancer-specific mortality rates. I estimated the cancer-specific mortality rates using the CODs obtained by the SRTR to compare with the cancer-attributable mortality rates estimated using PAFs. Based on SRTR CODs, I calculated a cancer-specific mortality rate of 368 per 100,000 person-years, substantially lower than the above cancer-attributable mortality rate (516 per 100,000 person-years). This difference is likely a reflection of inaccurate CODs. This may occur because recipients may appear to die from non-cancer CODs, such as infection or graft failure, that are downstream effects of cancer or its treatment. Although these estimates may reflect some inaccuracies, both estimates are vastly higher than the cancer mortality rate in the general population of 153 per 100,000 person-years.

In order to evaluate how many non-cancer CODs were likely due to cancer, I used the PAF methodology. With COD specified as cancer, a high proportion of deaths were attributed to cancer, as expected (PAF=74%). Notably, additional deaths were also attributed to cancer among deaths with other CODs. Specifically, I attributed to cancer 4.2% of deaths with infection recorded as the COD, 4.2% of deaths with graft failure as the COD, 5.6% of deaths with other recorded CODs, and 9.4% of unknown CODs.

The PAF methodology was chosen since it is likely that the CODs are indeed missing deaths caused by cancer. Determining single underlying CODs can be difficult because death may be precipitated by multiple causes and determination of the underlying COD is often subjective. This may be especially true for deaths in transplant recipients, who have multiple health conditions. The agreement of cancer vs. non-cancer COD from cancer registries compared to the SRTR was 70% with a kappa statistic of 0.4 indicating that COD may not be reliably recorded.

F. Life-years lost to cancer

This is the first study to estimate the LYL to cancer among recipients. The mean life-years lost per transplant recipient was 0.16 years within 10 years post-transplant and 2.7 years per cancer. This resulted in a loss due to cancer of 1.9% of all life-years that would be expected under a cancer-free scenario. The two cancer sites with the highest impact were lung cancer and NHL, which are common and associated with high mortality in transplant recipients.[83] These two cancer sites were responsible for 39% of the total LYL due to cancer in the cohort. Furthermore, once a recipient developed either lung cancer or NHL, they lost an average of almost five life-years.

Similar to the PAF results, the effect of immunosuppression on LYL is notable. Lung and heart recipients had the highest LYL to cancer. Moreover, lung cancer and NHL contributed a large portion of the total LYL in these recipients, especially among recipients age 50 and older. Lung and heart recipients receive the most intensive immunosuppression, which plays an important role in cancer incidence and mortality after transplantation.[18,83]

Among young recipients, age 35 and younger, the proportion of LYL was small (0.5%). However, NHL was a large contributor again indicating the role of immunosuppression. Among transplant recipients in this age group, NHL typically results from primary EBV infection.[68-70] It is important to note that the LYL that we estimated for young recipients were restricted by the censoring of follow-up imposed at 10 years after transplant. Indeed, the absolute LYL to cancer would be much larger for recipients if they were followed over a longer interval, and this would most substantially affect the estimates for younger recipients who have the longest potential lifetimes after transplant.

These results were derived from Cox regression models used to estimate the survival experience of the cohort overall and what would be expected under a cancer-free scenario. The Cox regression models under both scenarios were adjusted for the same observed confounders. The differences in the areas under the survival curves are not confounded since the same recipients comprise both groups and their baseline data are used to generate the survival curves. So there cannot be a variable that has a distribution in the overall cohort and different from the distribution in the cancer-free scenario since they are the same group at baseline. Moreover, I also estimated the LYL using Kaplan-Meier curves instead of Cox regression models. The results were very similar and not statistically different.

An assumption of any survival analysis is uninformative censoring. In the modeling of the cancer-free scenario, this assumption implies that the remaining recipients are representative of those who were censored at cancer (i.e. the censoring time is independent of the failure time) after controlling for covariates.

G. Life-years lost to cancer using a matching method

An alternative approach was to match recipients who developed cancer to those who did not to serve as a control group. Thus, the cancer-free scenario was comprised of the matched controls. The matching was conducted using the prognostic score which estimates the baseline risk of death using baseline information. Several approaches were used to estimate the survival experience and derive the LYL to cancer. They were Kaplan-Meier curves stratified by time since diagnosis/selection, Cox models with cancer diagnosis/selection as the delayed entry time, and Cox models starting the time scale at diagnosis/selection. I found that the total LYL to cancer was most similar from the estimates using the stratified Kaplan-Meier and the Cox models with time since diagnosis/selection as the time scale. This is likely because the risk sets generated using these methods are the most similar. The risk sets are the same except that the stratified Kaplan-Meier method uses stratification to limit the risk sets according to when the cancer occurred post-transplant whereas the Cox models using time from diagnosis/selection have a mix of when cancers occurred. The Cox models with delayed entry time resulted in higher estimates of the LYL to cancer. The risk set in this method is comprised of cancers and controls who have survived to time t since transplant. This difference is likely driven by an interaction between the time scales and when deaths occur. The Kaplan-Meier was selected to generate the final results using matching because it controls for when the cancer occurred post-transplant and the matching controls for differences in the baseline risk of death between cases and controls.

The total LYL estimates were higher using the modeling method (primary analyses presented in Chapter 0) compared to the matching method. Among kidney recipients, the total LYL was 38% higher using the modeling method compared to the matching method (19,148 (95% CI: 18,052, 19,983 vs. 13,875 (95% CI: 13,218, 14,565)). Similarly, the total LYL was 22% higher among liver recipients, 18% higher for heart recipients and 27% higher for recipients of other organs. Among lung recipients the total LYL was 51% higher using the modeling method, however, the absolute difference was small (2,878 LYL for modeling vs. 1,905 LYL for matching).

The matching method is conceptually straightforward and I was able to construct a control group that had good overlap on the prognostic score, and many matches were found for the vast majority of cancer cases. However, the modeling method was chosen because stratification of models was not needed, thereby providing a more streamlined method.

H. Strengths and limitations

Strengths of the studies in this dissertation include a large sample representative of the US transplant population, including recipients of all organ types. I obtained incident cancer information from population-based cancer registries that ascertain cancers in their catchment areas. Alternative statistical methods were utilized for both studies. Since two different approaches yielded similar results, I am confident that the estimates are consistent with the quantities we aimed to estimate.

This dissertation provides estimates of the cancer mortality burden that have not been previously published, such as the PAF and LYL to cancer. Previous studies have reported the SMR comparing the risk of cancer death among recipients to the general population.[2,5] Cancer mortality rates for kidney recipients have been previously published using COD data.[24] Kiberd et al. estimated a cancer mortality rate of 206 per 100,000 person-years among kidney recipients. This was fairly similar to the rate I found using COD of 297 per 100,000 person-years. However, my analysis showed that COD is often misclassified and thus the cancer-attributable mortality rate of 410 per 100,000 estimated from the PAF is more accurate. Furthermore, I was able to estimate cancer mortality rates for five transplanted organ types, population subgroups and from time since transplant.

A limitation is that cancer registries do not collect information on non-melanoma skin cancers. While these cancers cause substantial morbidity among transplant recipients,[56,57] they are not often fatal and, therefore, would not greatly contribute to mortality. Also, some cancers may have been missed by the cancer registries if the linkage did not match the cancer to the recipient, or if recipients migrated from the catchment area. However, outmigration is uncommon.[1]

I. Prevention and screening

There are opportunities to reduce cancer mortality in transplant recipients through prevention and screening. Most cancer screening guidelines for recipients follow recommendations for the general population, but these results indicate that more tailored guidelines would be useful.[48]

Given the high PAF and LYL for NHL, this cancer should be a priority. Since high levels of circulating EBV load can be used as a marker of PTLD, there may be opportunities to use this biomarker to screen for NHL, especially in pediatric recipients.[74] One practice guideline suggests that EBV serostatus be determined on all transplant recipients to identify those at high risk for PTLD and then monitored closely for clinical symptoms.[114] The baseline serostatus combined with routine EBV viral load monitoring can be used to detect EBV infection and estimate risk for PTLD.[115] This strategy could result in better prevention of PTLD and thus reduce risk of NHL.

Lung cancer is a major contributor to cancer mortality among recipients. Although high-risk current and heavy smokers in the general population are recommended for annual chest CT,[95] no recommendations are currently provided for recipients. Screening high risk recipients may provide a benefit. Lung recipients have unique considerations for screening. Among single lung recipients, lung cancers primarily arise in the native lung and the majority of recipients who develop lung cancer had COPD and idiopathic pulmonary fibrosis (IPF).[87] Recipients of single lung transplants may benefit from continuous screening for lung cancer especially since risk of lung cancer increases with time post-transplant.[87] Triplette et al. caution that interpretation of screening CT may be complicated by pre-existing lung disease and may not confer a survival benefit.

Liver cancer resulted in the highest LYL to cancer. Hepatitis C virus (HCV) infection is one of the main indications for liver transplantation, and these recipients are likely to experience re-infection with HCV in their donor liver. Treating underlying liver disease (e.g., with direct-acting antiviral medications for hepatitis C virus infection) either pre- or post-transplant may improve outcomes.[116] Screening for liver cancers in high-risk recipients should also be considered.[48,75]

Finally, cancer prevention strategies aimed at modifying lifestyle cancer risk factors, especially smoking cessation but also sun protection,[76] should be prioritized to reduce mortality from lung cancer and melanoma, respectively. Although melanoma was not a major contributor to cancer mortality, behavioral strategies aimed at sun protection may reduce incidence and mortality of skin cancer. Sun protection education may increase sun protection behaviors among recipients.[117,118]

J. Conclusion

In conclusion, this dissertation provides new population-based estimates of cancer mortality for transplant recipients in the United States. Cancer is a substantial cause of mortality among recipients and results in a shortened lifespan. Among recipients, the cancer-attributable mortality is increasing with age and time since transplant, indicating that deaths from cancer will increase in the future as survival following transplantation improves.

Appendix 1: Statistical considerations in estimating the population-attributable fraction

A. Modeling of age in the Cox regression model used to estimate the population-attributable fraction

Since age is strongly associated with cancer incidence and death, the functional form of how age should enter the Cox model was thoroughly examined. Several functions of age were examined in a model with cancer as a time-dependent variable.

Age in 5-year categories

In this model, age was entered as a categorical variable with 17 dummy categories resulting in estimating 16 parameters. The estimated beta coefficients were plotted (Figure 16). The plot shows that children age 0 to 4 have a higher mortality risk compared to children age 5 to 9. Children aged 5 to 19 have an increasing mortality risk that levels off starting at age 20 and then increases linearly with age starting at age 30.

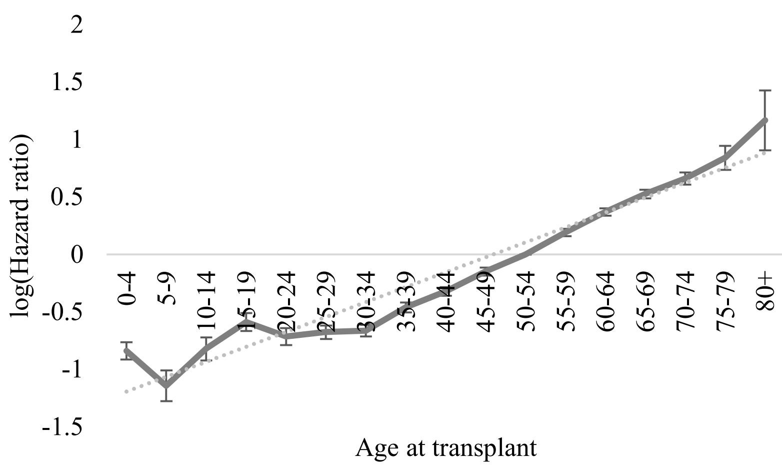


Figure 16: Estimated log-hazard ratios for age as a categorical variable from the Cox model used to estimate the population-attributable fraction

Age as restricted cubic spline

Age modeled with a restricted cubic spline can account for the non-linear effect of age on mortality risk. A restricted cubic spline was estimated using the method of Stone and Koo as implemented in the SAS macros written by Harrell.[119,120] This method fits a piecewise cubic polynomial with smooth transitions at the knots and restricted to linear ends. Figure 17 shows the estimated spline transformation for age. The knots, indicated by vertical lines, are chosen based on percentiles of the age distribution, the first knot is approximately at 13 years. The placement of this knot does not allow the spline to capture the increased mortality among the youngest ages.

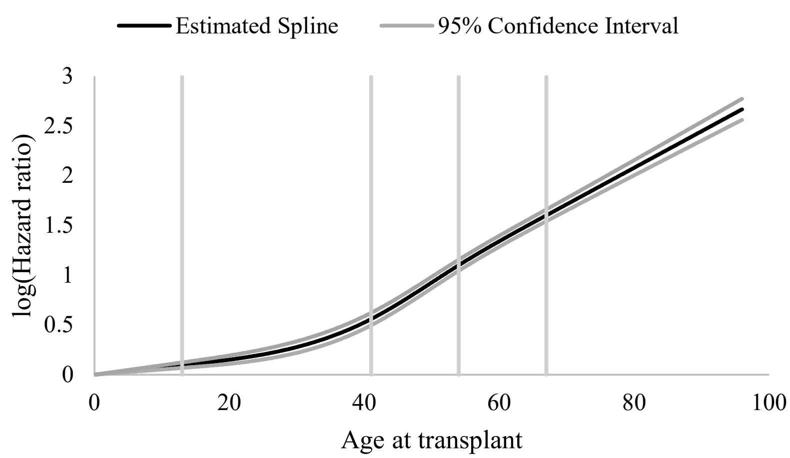


Figure 17: Estimated spline transformation of age from the Cox model used to estimate the population-attributable fraction

Knots are shown by vertical lines at age 13, 41, 54, and 67. Note: Only age was in the model that generated this plot.

Stratified by age

The baseline hazard of the Cox model was stratified by age in 5 strata (age 0-17, 18-34, 35-49, 50-64 and 65+).

Age as categorical within strata

In this model, age was categorical within each stratum. The categories were approximately 3- year age groups. The estimated log-hazard ratios were plotted to show the functional form for age within each strata. The oldest age group was selected as the reference group within each strata. The plot shows the same effect of age as when it is in 5 year categories. That is, mortality risk in the youngest stratum (light orange line) is highest among the 0 to 4 year olds and is not linear within this stratum (Figure 18). The other age stratum do show a linear effect of age.

Age as continuous within strata

In this model, the same strata were used but age was modeled as a continuous variable within each strata. Based on Figure 18, a linear effect for age is not appropriate for the youngest age stratum.

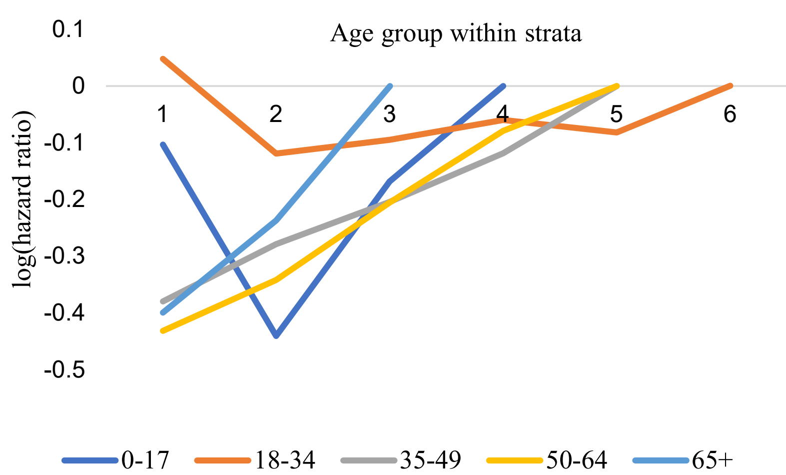


Figure 18: Estimated log-hazard ratios from Cox model stratified by age and then age as categorical within strata

Age modeled with a piecewise linear spline

Based on the model with age as 5-year categories, there was a changepoint in the mortality risk at about at 5 and 19. Age was entered in the model as a piecewise linear spline with knots chosen at age 5 and 19. The mortality risk was a linear slope in between the knots and the slopes could be different between the changepoints but were restricted so that the segments were joined at the knots. The log hazard ratio for age controlling for cancer incidence is shown in Figure 19.

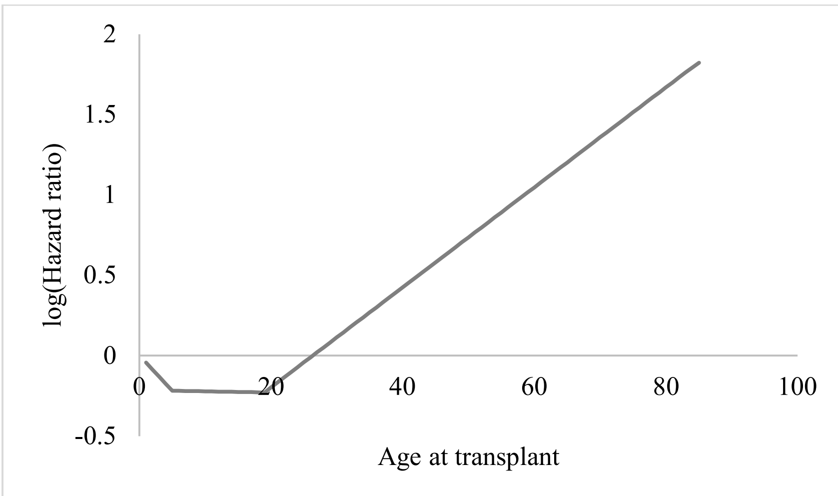


Figure 19: Estimated log hazard ratio association from model with age as a piecewise linear spline controlling for cancer

Summary

Table 32 shows the impact of the various functional forms of age used in the Cox regression models as measured by the AIC. There is not much difference in the AIC, however the spline models have far fewer degrees of freedom.

	DF	AIC*
Categorical Age (17 five-year age groups)	17	966637
Stratified on Age with Age as Categorical within Strata	19	966637
Stratified on Age with Age as Continuous within Strata	2	966734
Linear Piecewise Spline	4	1095281
Restricted Cubic Spline	4	1095150

DF=Degrees of freedom; AIC: Akaike information criterion
**Lower values of the statistic indicate a more desirable model

Table 32: Summary of AIC for Cox models with different functional forms for age

Impact on the estimate of cancer

Table 33 shows the estimates for cancer from the models described above with different functional forms for age. The estimate for cancer, which is the primary interest in this analysis, is very robust to the functional form for age.

	Estimate of cancer	**SE**	**Hazard Ratio**	**95% Confidence Interval**
Categorical Age (17 five-year age groups)	1.61	0.01	5.02	4.90, 5.15
Stratified on Age with Age as Categorical within Strata	1.60	0.01	4.94	4.81, 5.07
Stratified on Age with Age as Continuous within Strata	1.60	0.01	4.95	4.83, 5.08
Linear Piecewise Spline	1.62	0.01	5.04	4.91, 5.17
Restricted Cubic Spline	1.61	0.01	5.01	4.88, 5.14

Table 33: Estimated effect of cancer with different functional forms for age

The form chosen is age modeled as a linear piecewise spline. This was chosen since it captures the shape of the relationship and is more parsimonious than the model with 17 age groups.

B. Estimate of the population-attributable fraction in the presence of an interaction between cancer and follow-up time in the Cox model

Combining hazard ratios for cancer

A model was considered that included an interaction between cancer and the time since cancer diagnosis. This model is shown below with indicator variables that represent the interaction between incident cancer and time. Time since cancer diagnosis was categorized into four intervals, <1 year, 1-3 years, 3-5 years, and 5 or more years, thus four hazard ratios for cancer were estimated from the model.

This Cox model can be written as:

$$\lambda(t) = \beta_1 C_1 + \beta_3 C_3 + \beta_5 C_5 + \beta_{5+} C_{5+} + \gamma_1 Sex + \gamma_2 Race + \gamma_3 Year + \gamma_5 Age \quad (1)$$

where: t is the time since transplant

$$C_1 = \begin{cases} 1 \text{ if person has cancer and it is 1 year following cancer diagnosis} \\ 0 \text{ otherwise} \end{cases}$$

$$C_3 = \begin{cases} 1 \text{ if person has cancer and it is more than 1 year but less than 3 years following cancer diagnosis} \\ 0 \text{ otherwise} \end{cases}$$

$$C_5 = \begin{cases} 1 \text{ if person has cancer and it is more than 3 years but less than 5 years following cancer diagnosis} \\ 0 \text{ otherwise} \end{cases}$$

$$C_{5+} = \begin{cases} 1 \text{ if person has cancer and it is more than 5 years following cancer diagnosis} \\ 0 \text{ otherwise} \end{cases}$$

In order to obtain one PAF to represent the entire follow-up period, these hazard ratios must be combined.

Extension of formula using one hazard ratio

The PAF associated with cancer can be calculated by the following:

$$PAF = p_d \times \left(\frac{HR-1}{HR}\right) \quad (2)$$

where p_d is the percentage of deaths with a prior cancer diagnosis among all deaths in the cohort and HR is the estimated hazard ratio for cancer.

An extension of this formula is to calculate the PAF associated with cancer for each time interval and then to sum over each time interval to obtain the PAF for the entire follow-up. Let D_{total} be the number of deaths in the cohort and D_{cancer} be the number of deaths among those with cancer in the cohort. Then the percentage of deaths with a prior cancer diagnosis in a particular time interval is:

$p_{d_i} = \frac{D_{cancer,i}}{D_{total}}$ where i= <1, 1-2.99, 3, 4.99, 5+ years between cancer and death (3)

Then the PAF in each interval is calculated by:

$$PAF_i = p_{d_i} * \left(\frac{HR_i - 1}{HR_i}\right) \tag{4}$$

And then summed over each time interval:

$$PAF = \sum_{i=1}^{4} PAF_i \tag{5}$$

No variance estimator is readily available for this PAF quantity.

Attributable risk methods for survival analysis

Given that the outcome under study was time-to-event and an interaction between cancer and time since cancer diagnosis was identified, attributable risk methods for survival analysis were considered. Gassama et al.[121] reviewed several methods for estimating the attributable risk in this setting. Four methods were reviewed and tested in a simulation study. Three of these methods define the attributable risk as a function of time and so there is not one summary measure of the attributable risk over the whole follow-up period. However, the method proposed by Spiegelman et al.[122] does produce one single estimate of the attributable risk in a cohort study. In this method, the relative risk of cancer can be estimated using the Cox model and the person-years of follow-up are used for the exposure prevalence in the cohort.

The PAF is then estimated by the following extension of the PAF for multivariate exposures:

$$PAF = \frac{\sum_{s=1}^{S} p_s(RR_s - 1)}{1 + \sum_{s=1}^{S} p_s(RR_s - 1)} = 1 - \frac{1}{\sum_{s=1}^{S} p_s(RR_s)} \qquad (6)$$

In this model, the levels of exposure are the stratum created by the cancer by time since diagnosis interaction. This is a benefit since it accounts for the person-time accumulated in each cancer by time since diagnosis interaction. The variance of the PAF was derived and is implemented by the authors in a SAS macro.

This method performed well in the simulation study with relatively no bias and good coverage probabilities.[121] However, the simulation tested this method using an unadjusted binary exposure and with proportional hazards. A limitation of this study is that the PAF adjusted for covariates was not evaluated.

Discussion

Both extensions of the PAF given in equations 2 and 6 are valid if the exposure of interest is not confounded. Indeed, the simulation study showed good performance of the estimator in equation 6 under ideal conditions (binary exposure, proportional hazards, unadjusted relative risks). However, the PAF in equation 6 is not valid if confounding exists.[28,123] Rather, the multivariate extension of equation 2 should be used and is given in equation 5.

A challenge to applying this formula to the model with an interaction between cancer and time since diagnosis is that a variance estimator is not readily available. The difference between estimates for cancer given in equation 1 and the simpler model with cancer as a time-dependent variable is in how the risk of death from cancer is averaged over the follow-up period. Specifically, in the model with an interaction with time, four estimates of the risk of death are produced and then averaged in the calculation of the PAF weighted by person-time. In the simpler model, the risk of death from cancer is averaged over follow-up time on the hazard scale and this summary risk is used in the calculation of the PAF.

Table 34 shows the estimates from each of these methods: Method 1 uses hazard ratios estimated from the simple model with cancer as a time-dependent variable; Method 2 uses the four hazard ratios estimated from Model 1 and combined using equation 5; Method 3 uses the four hazard ratios estimated from Model 1 and combined using equation 6.

The PAF estimates from Method 1 and 2 were similar using both adjusted and unadjusted hazard ratios. The PAF estimated from Method 3 were similar but only using the unadjusted hazard ratios. Since the PAF formulation used in Method 3 is valid for unadjusted hazard ratios, the results were expected to be similar. The adjusted PAF estimates using Method 3 were consistently lower than the PAF estimates produced using Methods 1 and 2. This is likely because the PAF formulation is not appropriate for use with adjusted risk estimates.

In conclusion, since the PAF point estimates for Method 1 and 2 are almost equal in each stratum and the variance estimate using Method 2 is not available, Method 1 was used to compute the PAF estimates in the final analysis as shown in chapter 4.

	Unadjusted					Adjusted				
			Model incl interaction with time					Model incl interaction with time		
	Method 1*		Method 2†	Method 3‡		Method 1*		Method 2†	Method 3‡	
	PAF	95% CI‖	PAF	PAF	95% CI	PAF	95% CI‖	PAF	PAF	95% CI
Overall	0.139	(0.132, 0.147)	0.138	0.139	(0.131, 0.146)	0.132	(0.124, 0.139)	0.13	0.105	(0.099, 0.111)
Sex										
Female	0.117	(0.105, 0.128)	0.116	0.117	(0.105, 0.128)	0.111	(0.100, 0.122)	0.11	0.092	(0.082, 0.102)
Male	0.153	(0.143, 0.162)	0.151	0.151	(0.142 0.161)	0.144	(0.134, 0.153)	0.141	0.113	(0.105, 0.121)
Race										
White	0.15	(0.141, 0.158)	0.148	0.149	(0.14, 0.159)	0.142	(0.134, 0.151)	0.14	0.117	(0.109, 0.125)
Black	0.109	(0.092, 0.126)	0.108	0.108	(0.091, 0.125)	0.102	(0.085, 0.118)	0.1	0.08	(0.066, 0.094)
Hispanic	0.112	(0.088, 0.136)	0.111	0.113	(0.091, 0.134)	0.107	(0.083, 0.130)	0.106	0.083	(0.066, 0.101)
API	0.143	(0.095, 0.192)	0.142	0.147	(0.105, 0.188)	0.138	(0.090, 0.186)	0.137	0.11	(0.076, 0.144)
Other	0.128	(0.038, 0.219)	0.127	0.119	(0.026, 0.21)	0.124	(0.034, 0.214)	0.123	0.095	(0.015, 0.174)
Organ										
Kidney	0.14	(0.128, 0.152)	0.139	0.134	(0.123, 0.144)	0.131	(0.119, 0.143)	0.129	0.094	(0.086, 0.102)
Liver	0.134	(0.120, 0.149)	0.133	0.144	(0.126, 0.161)	0.13	(0.116, 0.145)	0.129	0.121	(0.105, 0.137)
Heart	0.166	(0.150, 0.181)	0.163	0.164	(0.144, 0.183)	0.164	(0.148, 0.179)	0.161	0.156	(0.137, 0.175)
Lung	0.106	(0.093, 0.119)	0.105	0.109	(0.085, 0.133)	0.105	(0.092, 0.118)	0.104	0.104	(0.08, 0.127)
Other	0.099	(0.071, 0.127)	0.098	0.102	(0.072, 0.132)	0.098	(0.070, 0.126)	0.097	0.097	(0.068, 0.126)

Table 34: Comparison of PAF estimates from 3 methods

Table 34: Continued

	Unadjusted					Adjusted				
			Model incl interaction with time					Model incl interaction with time		
	Method 1*		Method 2†	Method 3‡		Method 1*		Method 2†	Method 3‡	
	PAF	95% CI‖	PAF	PAF	95% CI	PAF	95% CI‖	PAF	PAF	95% CI
Transplant Year										
1987-1994	0.156	(0.144, 0.168)	0.154	0.153	(0.137, 0.169)	0.145	(0.133, 0.157)	0.143	0.113	(0.1, 0.126)
1995-1999	0.14	(0.129, 0.151)	0.139	0.138	(0.125, 0.151)	0.13	(0.120, 0.141)	0.129	0.102	(0.091, 0.113)
2000-2004	0.136	(0.123, 0.150)	0.135	0.139	(0.124, 0.154)	0.13	(0.116, 0.143)	0.129	0.106	(0.094, 0.119)
2005-2008	0.123	(0.099, 0.147)	0.122	0.13	(0.11, 0.155)	0.119	(0.095, 0.143)	0.118	0.104	(0.085, 0.123)
2009-2014	0.095	(0.026, 0.164)	0.095	0.109	(0.068, 0.149)	0.094	(0.025, 0.162)	0.093	0.088	(0.054, 0.122)
Age										
0-17	0.063	(0.026, 0.100)	0.063	0.065	(0.032, 0.098)	0.061	(0.025, 0.098)	0.061	0.053	(0.024, 0.082)
18-34	0.075	(0.050, 0.099)	0.074	0.074	(0.054, 0.094)	0.073	(0.049, 0.097)	0.073	0.065	(0.047, 0.083)
35-49	0.106	(0.093, 0.120)	0.105	0.101	(0.088, 0.114)	0.105	(0.091, 0.119)	0.104	0.095	(0.082, 0.107)
50-64	0.158	(0.148, 0.168)	0.155	0.154	(0.142, 0.165)	0.155	(0.145, 0.165)	0.152	0.141	(0.131, 0.152)
65+	0.156	(0.140, 0.173)	0.154	0.152	(0.131, 0.172)	0.155	(0.138, 0.171)	0.152	0.144	(0.124, 0.164)

*One HR for cancer estimated from model. PAF computed using p_d((HR-1)/HR), variance estimated using method of Engels et al. (2017)

†Model includes interaction with time since cancer diagnosis, so 4 HRs for cancer were estimated. PAF computed using pd((HR-1)/HR) for each time period and then summed. 95% CI formula is N/A. ‡Model includes interaction with time since cancer diagnosis, so 4 HRs for cancer were estimated. PAF computed using method of Spiegleman et al. (2007). ‖95% CI calculated using +/- 1.96*sqrt(var(PAF))

C. Validity of proportional hazards assumption for the Cox model used to estimate the population-attributable fraction

Validity of proportional hazards assumption of other covariates for aim 1.1

The proportional hazards assumption for each covariate was evaluated by plotting the survival probability (S(t)) vs. time (t) and the log(-log(S(t)) vs. log(t). The survival probability was estimated using the Kaplan-Meier and so it is the unadjusted probability of survival. The plot of S(t) vs. time may indicate non-proportional hazards if the survival curves are crossing. Moreover, if the lines for each stratum in the plot of log(-log(S(t)) vs. log(t) are parallel then that indicates that the assumption of proportional hazards has been met.

Figure 20 to Figure 23 show these plots for each covariate of interest. By examination of the plots, race, sex, and year of transplant all meet the assumption of proportional hazards. The plots by transplanted organ type may indicate non-proportional hazards. In the plot of log(-log(S(t)) vs. log(t), the line for other/multi organs crosses several of the other lines. Similarly, the estimate for pancreas crosses the estimates for the other organs. Since the combination organ groups are small, they were combined into the Other/Multi category. The plots indicate that the proportional hazards assumption is met (Figure 24).

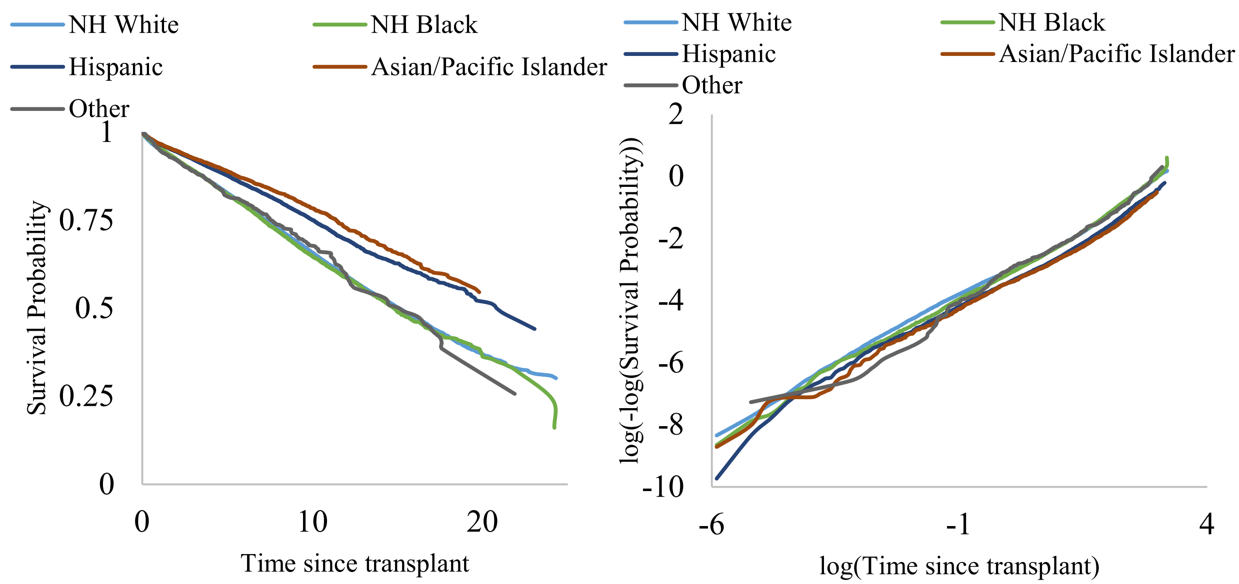


Figure 20: Survival plots by race

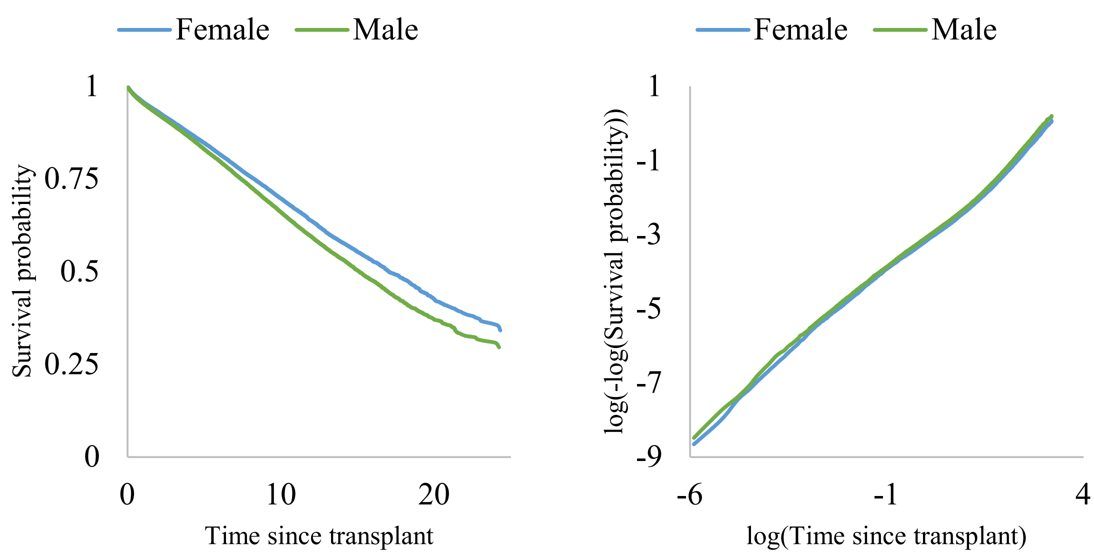


Figure 21: Survival plots by sex

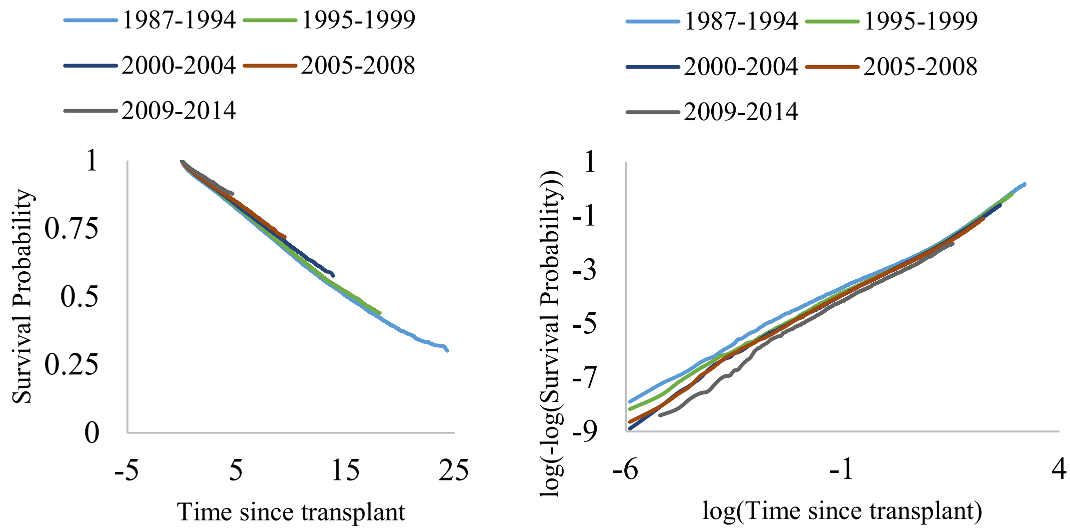


Figure 22: Survival plots by year of transplant

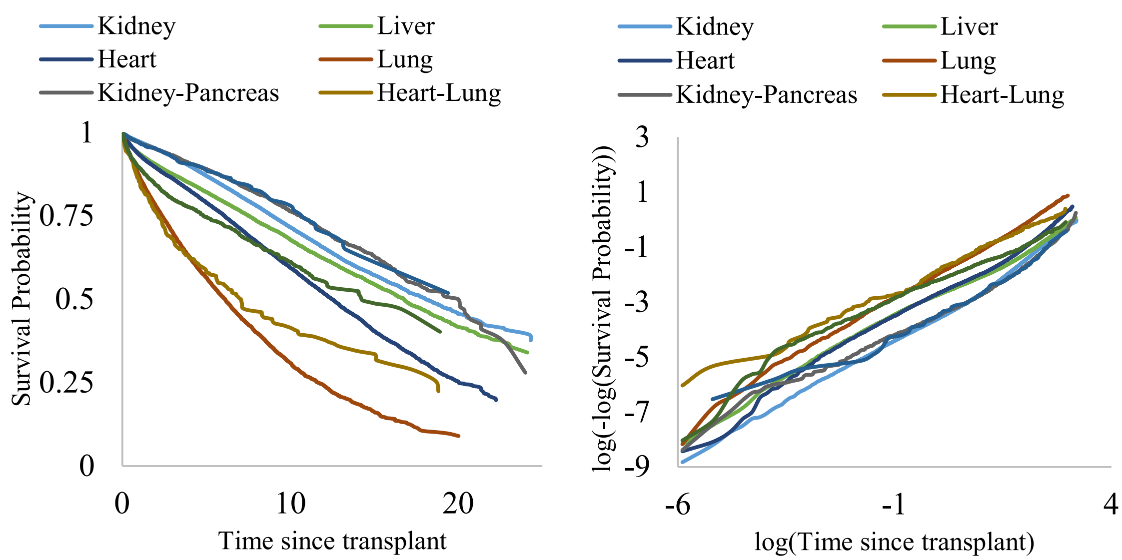


Figure 23: Survival plots by transplanted organ

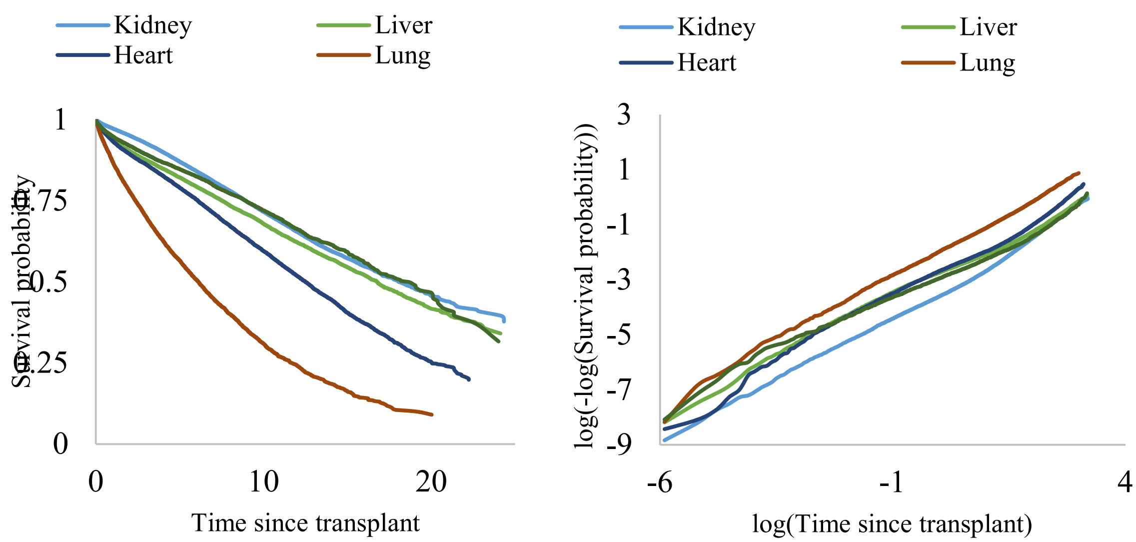


Figure 24: Survival plots by transplanted organ with collapsed groupings

Adjusted model

The model was adjusted for sex, age, year of transplant, race and organ transplant. Cancer was estimated as a time-dependent covariate.

Schoenfeld residuals

The residuals were estimated from the adjusted model. These were plotted against time to evaluate non-proportional hazards for each variable (Figure 25 to Figure 28). A pattern in the residuals indicates that the proportional hazards assumption may not be met.

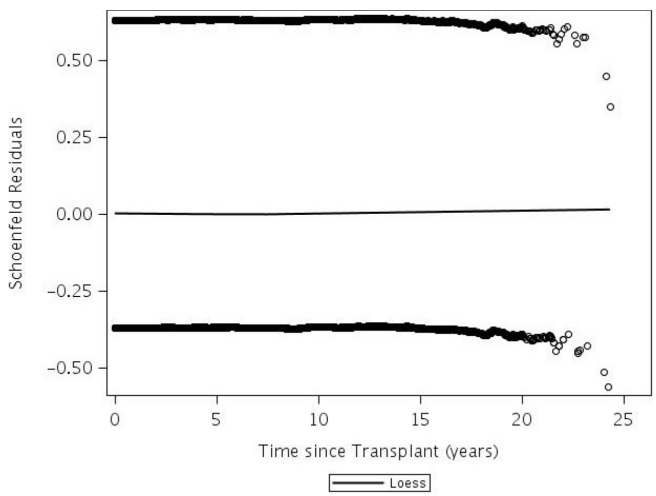


Figure 25: Schoenfeld residuals with loess for sex

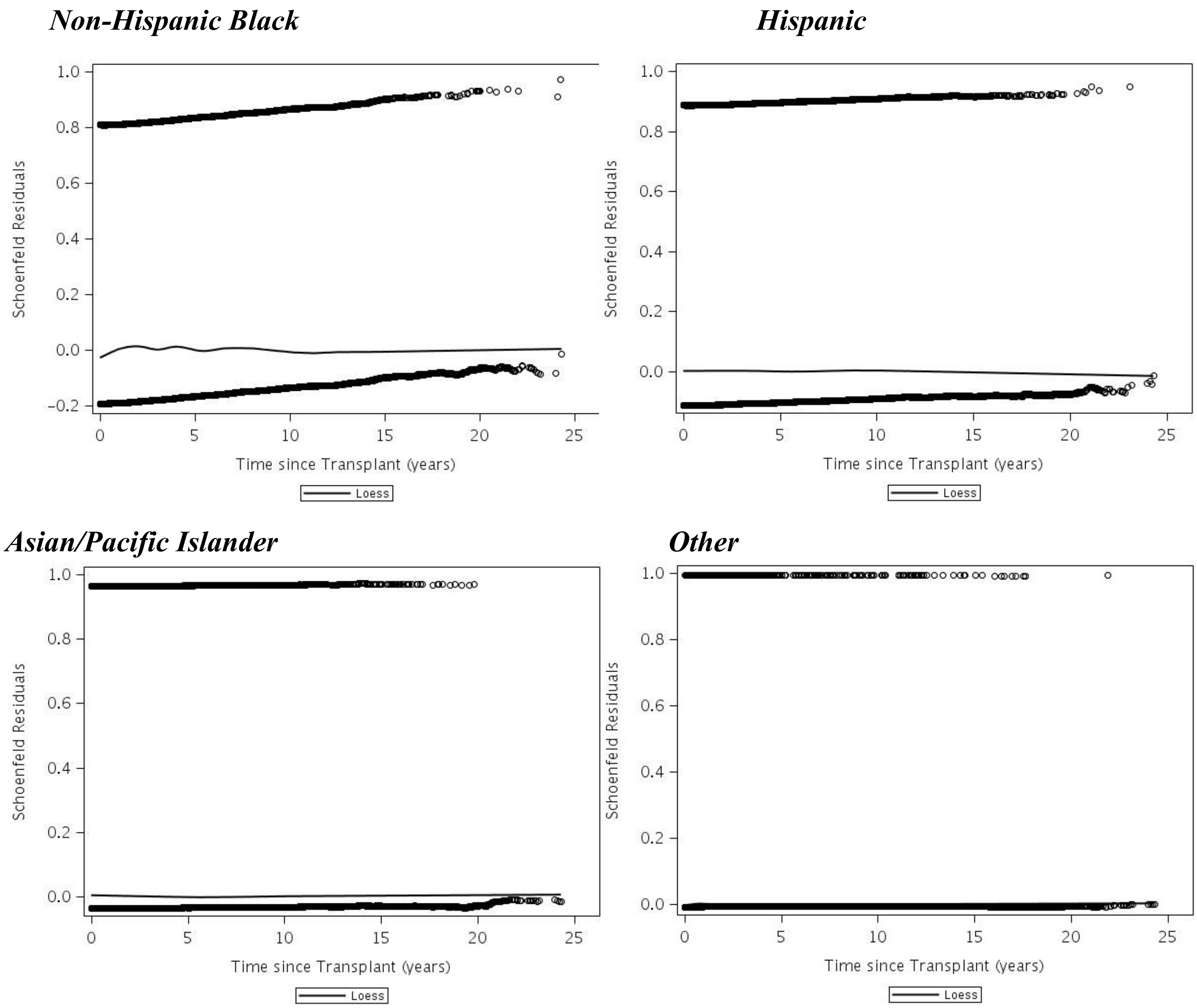


Figure 26: Schoenfeld residuals with loess by race (Non-Hispanic white is reference group)

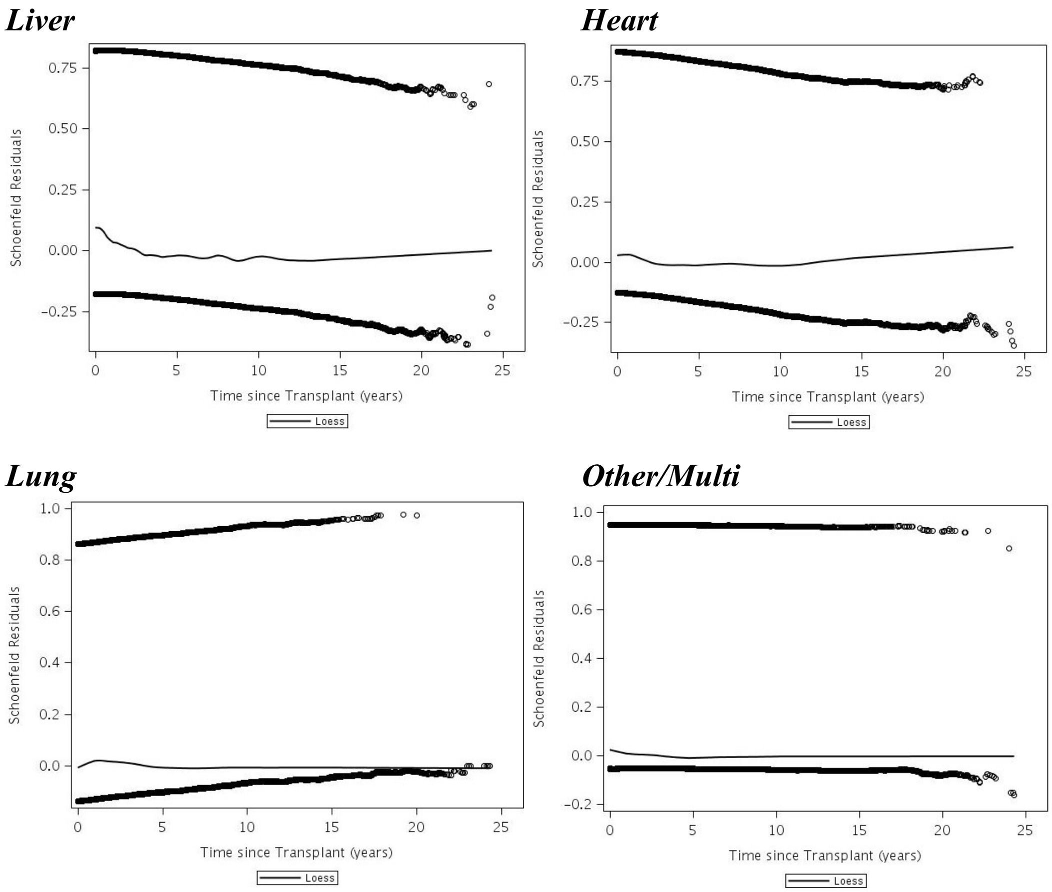


Figure 27: Schoenfeld residuals with loess by organ (kidney is reference group)

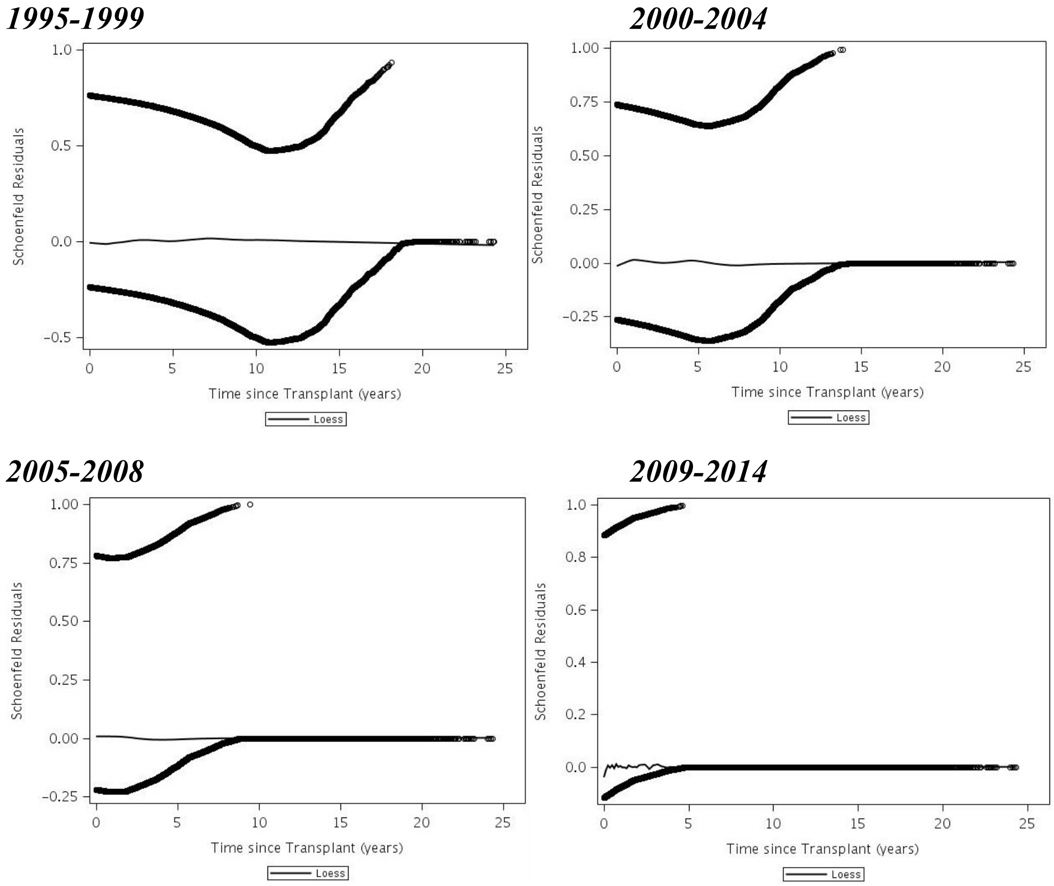


Figure 28: Schoenfeld residuals with loess by year of transplant (1987-1994 is reference group)

Assessment of residuals

The residual plots for sex, race and organ all show a linear trend without a definitive pattern. However, the plots for year at transplant show a definite pattern indicating an interaction with this variable.

Separate models by sex, race, organ and year of transplant

Separate models were run to assess if there were any interactions between the covariate and cancer. The models are shown in Table 35 to Table 38.

Sex		Women n=86,734		Men n=135,228	
		HR	**95% CI**	**HR**	**95% CI**
Cancer		4.79	(4.58, 5.00)	4.79	(4.64, 4.94)
Race					
	NH white	Ref			
	NH black	1.30	(1.25, 1.35)	1.28	(1.24, 1.32)
	Hispanic	0.94	(0.90, 0.99)	0.93	(0.90, 0.97)
	API	0.72	(0.66, 0.78)	0.85	(0.79, 0.90)
	Other	1.09	(0.91, 1.32)	1.20	(1.04, 1.39)
Year of transplant					
	1987-1994	Ref		Ref	
	1995-1999	0.89	(0.86, 0.93)	0.85	(0.82, 0.88)
	2000-2004	0.78	(0.75, 0.82)	0.74	(0.72, 0.77)
	2005-2008	0.66	(0.63, 0.70)	0.66	(0.64, 0.69)
	2009-2014	0.55	(0.50, 0.60)	0.54	(0.50, 0.57)
Organ Transplanted					
	Kidney	Ref		Ref	
	Liver	1.22	(1.18, 1.27)	1.27	(1.23, 1.31)
	Heart	1.70	(1.62, 1.79)	1.25	(1.21, 1.29)
	Lung	3.74	(3.57, 3.92)	3.18	(3.05, 3.32)
	Other/Multi	1.58	(1.48, 1.68)	1.32	(1.25, 1.39)
Age at transplant; 1 year increase					
	0-5	1.01	(0.97, 1.05)	0.96	(0.92, 1.00)
	6-19	0.99	(0.94, 1.04)	1.04	(0.99, 1.09)
	20+	1.04	(1.03, 1.05)	1.04	(1.03, 1.05)

HR=hazard ratio, CI=confidence interval, NH=non-Hispanic, API=Asian/Pacific islander

Table 35: Relative risk of death among recipients by sex

Table 36: Relative risk of death among recipients by transplanted organ

Organ transplanted	Kidney		Liver		Heart		Lung		Other/Multi	
	n=136,534		n=39,155		n=22,912		n=10,467		n=12,894	
	HR	**95% CI**	**HR**	**95% CI**	**HR**	**95% CI**	**HR**	**95% CI**	**HR**	**95% CI**
Cancer	4.67	(4.50, 4.85)	5.63	(5.32, 5.97)	4.6	(4.34, 4.88)	4.04	(3.71, 4.39)	5.8	(5.12, 6.57)
Sex										
Female vs. N	0.9	(0.88, 0.93)	0.89	(0.85, 0.93)	1.02	(0.97, 1.08)	0.99	(0.94, 1.05)	1.07	(0.99, 1.15)
Race										
NH white	Ref		Ref		Ref		Ref		Ref	
NH black	1.23	(1.19, 1.27)	1.41	(1.32, 1.52)	1.57	(1.48, 1.67)	1.03	(0.92, 1.16)	1.11	(0.98, 1.26)
Hispanic	0.88	(0.84, 0.91)	0.98	(0.92, 1.04)	1.1	(1.01, 1.20)	1.15	(1.01, 1.30)	0.95	(0.83, 1.09)
API	0.73	(0.68, 0.77)	0.81	(0.72, 0.91)	1.05	(0.91, 1.22)	0.93	(0.71, 1.21)	1.14	(0.87, 1.50)
Other	1.18	(1.00, 1.39)	1	(0.77, 1.30)	1.17	(0.85, 1.62)	1.46	(0.99, 2.15)	1.07	(0.76, 1.51)
Year of transplant										
1987-1994	Ref		Ref		Ref		Ref		Ref	
1995-1999	0.81	(0.78, 0.84)	0.94	(0.89, 0.99)	0.87	(0.82, 0.92)	0.93	(0.85, 1.02)	0.85	(0.76, 0.94)
2000-2004	0.67	(0.65, 0.70)	0.89	(0.84, 0.94)	0.82	(0.77, 0.87)	0.74	(0.68, 0.81)	0.68	(0.60, 0.77)
2005-2008	0.52	(0.50, 0.55)	0.94	(0.87, 1.01)	0.73	(0.67, 0.80)	0.71	(0.64, 0.78)	0.62	(0.54, 0.72)
2009-2014	0.4	(0.36, 0.43)	0.79	(0.70, 0.89)	0.6	(0.51, 0.70)	0.64	(0.56, 0.73)	0.6	(0.49, 0.75)
Age at transplant; 1										
0-5	0.89	(0.79, 0.99)	0.97	(0.92, 1.04)	1.01	(0.97, 1.06)	1.16	(1.03, 1.30)	1.01	(0.93, 1.09)
5-19	1.17	(1.03, 1.33)	1.1	(1.01, 1.19)	1.01	(0.95, 1.07)	0.83	(0.73, 0.94)	0.9	(0.81, 0.99)
20+	1.01	(0.99, 1.03)	0.95	(0.94, 0.97)	0.98	(0.97, 1.00)	1.05	(1.03, 1.08)	1.13	(1.10, 1.15)

HR=Hazard ratio, CI=confidence interval, NH=non-Hispanic, API=Asian/Pacific islander

Table 37: Relative risk of death among recipients by age at transplant

Age at transplant	0-17 n=16707			18-34 n=34997			35-49 n=67595			50-64 n=82092			65+ n=20571	
	HR	95% CI		HR	95% CI		HR	95% CI		HR	95% CI		HR	95% CI
Cancer	5.56	(4.65, 6.65)		5.59	(5.00, 6.24)		5.37	(4.82, 5.37)		4.7	(4.54, 4.87)		4.16	(3.91, 4.43)
Sex														
Female vs. Male	1.01	(0.92, 1.11)		1.1	(1.04, 1.17)		0.93	(0.86, 0.93)		0.93	(0.90, 0.95)		0.89	(0.84, 0.93)
Race														
NH White	Ref			Ref			Ref			Ref			Ref	
NH black	1.79	(1.59, 2.02)		1.41	(1.30, 1.52)		1.38	(1.25, 1.38)		1.25	(1.20, 1.30)		1.12	(1.04, 1.20)
Hispanic	1.18	(1.04, 1.33)		0.81	(0.73, 0.89)		0.9	(0.80, 0.90)		0.99	(0.95, 1.04)		0.95	(0.88, 1.04)
API	1.05	(0.80, 1.37)		0.72	(0.61, 0.86)		0.7	(0.57, 0.70)		0.86	(0.80, 0.92)		0.82	(0.73, 0.92)
Other	1.32	(0.86, 2.02)		0.95	(0.62, 1.45)		1.69	(1.11, 1.69)		1.01	(0.85, 1.20)		1.35	(0.97, 1.87)
Year of transplant														
1987-1994	Ref			Ref			Ref			Ref			Ref	
1995-1999	0.92	(0.81, 1.05)		0.84	(0.78, 0.91)		0.86	(0.79, 0.86)		0.88	(0.85, 0.92)		0.9	(0.83, 0.98)
2000-2004	0.77	(0.67, 0.88)		0.72	(0.66, 0.78)		0.73	(0.66, 0.73)		0.76	(0.73, 0.79)		0.83	(0.77, 0.90)
2005-2008	0.61	(0.52, 0.72)		0.65	(0.58, 0.72)		0.63	(0.55, 0.63)		0.67	(0.64, 0.71)		0.7	(0.64, 0.77)
2009-2014	0.5	(0.38, 0.66)		0.49	(0.40, 0.61)		0.53	(0.40, 0.53)		0.56	(0.52, 0.61)		0.58	(0.51, 0.66)
Organ tran														
Kidney	Ref			Ref			Ref			Ref			Ref	
Liver	2.13	(1.81, 2.51)		2.19	(1.98, 2.42)		1.7	(1.56, 1.70)		1.09	(1.06, 1.13)		0.86	(0.80, 0.92)
Heart	6.78	(5.88, 7.82)		4.24	(3.88, 4.63)		1.89	(1.69, 1.89)		1.07	(1.03, 1.11)		0.73	(0.68, 0.79)
Lung	18.85	(15.62, 22.75)		8.1	(7.37, 8.91)		4.36	(3.80, 4.36)		2.85	(2.73, 2.98)		2.3	(2.09, 2.54)
Other/Multi	12.93	(10.78, 15.51)		2.21	(2.02, 2.43)		1.48	(1.31, 1.48)		1.19	(1.09, 1.29)		1.27	(0.98, 1.64)
Age														
0-4	0.65	(0.57, 0.74)	18-20	0.99	(0.89, 1.12)	36-37	0.76	(0.68, 0.76)	50-52	0.65	(0.62, 0.68)	65-69	0.65	(0.57, 0.73)
5-9	0.59	(0.50, 0.70)	21-23	0.85	(0.76, 0.95)	38-40	0.83	(0.74, 0.83)	53-55	0.71	(0.68, 0.74)	70-75	0.77	(0.68, 0.87)
10-14	0.8	(0.69, 0.91)	24-26	0.92	(0.83, 1.02)	41-43	0.88	(0.80, 0.88)	56-58	0.81	(0.77, 0.84)	76+	Ref	
15-17	Ref		27-29	0.95	(0.87, 1.05)	44-46	0.94	(0.85, 0.94)	59-61	0.91	(0.87, 0.95)			
			30-32	0.92	(0.84, 1.01)	47-49	Ref		62-64	Ref				
			33-34	Ref										

HR=hazard ratio, CI=confidence interval, NH=non-Hispanic, API=Asian/Pacific islander

Table 38: Relative risk of death among recipients by transplant year

Year of transplant		1987-1994 n=32,257		1995-1999 n =49,140		2000-2004 n =58,297		2005-2008 n =50,901		2009-2014 n =31,367	
		HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Cancer		4.12	(3.93, 4.33)	4.15	(3.97, 4.34)	5.37	(5.11, 5.64)	7.48	(6.94, 8.06)	10.67	(8.99, 12.65)
Sex											
	Female vs.	0.93	(0.90, 0.97)	0.96	(0.93, 0.99)	0.96	(0.93, 1.00)	0.92	(0.87, 0.97)	0.96	(0.86, 1.06)
Race											
	NH white										
	NH black	1.31	(1.24, 1.38)	1.27	(1.22, 1.33)	1.3	(1.24, 1.36)	1.35	(1.26, 1.44)	1.33	(1.16, 1.53)
	Hispanic	0.87	(0.81, 0.94)	0.92	(0.87, 0.97)	0.96	(0.91, 1.02)	1.01	(0.94, 1.09)	0.99	(0.85, 1.16)
	API	0.87	(0.78, 0.96)	0.7	(0.63, 0.77)	0.8	(0.73, 0.88)	0.88	(0.78, 1.00)	0.84	(0.66, 1.06)
	Other	0.83	(0.65, 1.06)	1.47	(1.21, 1.80)	1.14	(0.90, 1.45)	1.25	(0.96, 1.63)	1.63	(0.96, 2.77)
Organ transplanted											
	Kidney										
	Liver	1.03	(0.98, 1.08)	1.11	(1.07, 1.16)	1.29	(1.23, 1.35)	1.95	(1.82, 2.08)	2.62	(2.28, 3.01)
	Heart	1.3	(1.23, 1.36)	1.23	(1.17, 1.28)	1.39	(1.32, 1.46)	1.73	(1.60, 1.88)	2.25	(1.90, 2.68)
	Lung	3.31	(3.06, 3.58)	3.44	(3.25, 3.64)	3.16	(2.98, 3.35)	4.01	(3.73, 4.32)	5.06	(4.39, 5.83)
	Other/Mul	1.36	(1.24, 1.49)	1.35	(1.25, 1.44)	1.35	(1.24, 1.46)	1.7	(1.52, 1.90)	2.59	(2.12, 3.17)
Age at transplant; 1											
	0-5	1.01	(0.96, 1.07)	1.04	(0.99, 1.10)	0.97	(0.92, 1.03)	0.84	(0.77, 0.92)	0.82	(0.69, 0.97)
	5-19	1	(0.93, 1.07)	0.94	(0.89, 1.00)	1.02	(0.95, 1.09)	1.2	(1.07, 1.34)	1.24	(1.00, 1.54)
	20+	1.02	(1.01, 1.04)	1.06	(1.04, 1.07)	1.05	(1.03, 1.07)	1.02	(1.00, 1.05)	1.02	(0.97, 1.08)

HR=hazard ratio, CI=confidence interval, NH=non-Hispanic; API=Asian/Pacific islander

Assessment of models

The point estimates of the hazard ratio for cancer are the same for men and women (Table 38). There are some differences in the hazard ratio for cancer by organ and age. However, the hazard ratio for cancer varies widely by year of transplant. Specifically, the hazard ratio is 4.12 and 4.15 for 1987-1994 and 1995-1999 respectively. The hazard ratio increases to 5.37 for transplant year 2000-2004 and is the highest in the most recent period at 10.67.

One hypothesis is that there is an interaction between transplant year and duration of follow-up. Transplants in 2009-2014 had much shorter follow-up time to assess cancer and mortality.

Interaction between transplant year and duration of follow-up

Model 1: Censor all recipients 3 years post-transplant

In this model, all recipients are censored 3 years post-transplant thus giving all recipients equal follow-up time for cancer and mortality assessment. If the mortality risk is the same for each year of transplant, then there is evidence of an interaction between transplant year and duration of follow-up.

Table 39 shows separate models by transplant year. The estimates of the hazard ratio for death for cancer are very similar which supports the hypothesis that there is an interaction between transplant year and duration of follow-up.

Model 2: Model interaction between cancer and time since cancer

In this model, the hazard ratio for death was allowed to vary by the time since cancer diagnosis. To explore the hypothesis that mortality associated with cancer is the highest soon after the cancer diagnosis models were fit allowing the hazard ratio for death to vary by the time since cancer diagnosis. Table 40 shows that the hazard ratio for death does vary by time since cancer diagnosis. Specifically, the risk of death is highest soon after a cancer diagnosis.

Since the hazards ratios for the categories of time since cancer diagnosis were similar, these were combined to simplify the model (Table 41).

Next this effect of cancer was evaluated by testing the interaction between cancer and follow-up time. Since risk of death varies by time since cancer diagnosis, this interaction has several terms to represent the effect of cancer changing over the follow-up (Table 42). Although the interaction terms are statistically significant, the estimated coefficients are small (range: -0.034 to -0.013). This interaction effect was eliminated from the final model since the interaction does not have a substantial impact on the resulting hazard ratios and to retain a more parsimonious model. Thus, the model shown in Table 41 is the final adjusted model.

Transplant year	1987-1999		1995-1999		2000-2004		2005-2008		2009-2014	
	HR	**95% CI**	**HR**	**95% CI**	**HR**	**95% CI**	**HR**	**95% CI**	**HR**	**95% CI**
Cancer	11.79	(10.53, 13.19)	9.32	(8.45, 10.27)	9.34	(8.55, 10.20)	8.88	(8.03, 9.82)	10.82	(9.00, 13.01)
Sex										
Female vs. Male	0.88	(0.82, 0.95)	0.99	(0.94, 1.05)	0.98	(0.93, 1.04)	0.92	(0.86, 0.98)	1	(0.90, 1.11)
Race										
NH White	Ref		Ref		Ref		Ref		Ref	
NH black	1.43	(1.30, 1.58)	1.38	(1.27, 1.49)	1.43	(1.34, 1.54)	1.39	(1.28, 1.50)	1.31	(1.14, 1.52)
Hispanic	0.93	(0.82, 1.06)	0.97	(0.88, 1.06)	0.99	(0.91, 1.07)	1	(0.92, 1.10)	1.02	(0.86, 1.19)
API	1.25	(1.05, 1.48)	0.82	(0.69, 0.97)	0.88	(0.76, 1.01)	0.97	(0.84, 1.13)	0.87	(0.68, 1.11)
Other	0.68	(0.39, 1.17)	1.58	(1.12, 2.21)	1.29	(0.92, 1.82)	1.23	(0.90, 1.67)	1.57	(0.91, 2.71)
Organ transplanted										
Kidney	Ref		Ref		Ref		Ref		Ref	
Liver	1.46	(1.33, 1.61)	1.58	(1.47, 1.71)	1.76	(1.64, 1.88)	2.37	(2.20, 2.57)	2.78	(2.41, 3.21)
Heart	1.72	(1.56, 1.89)	1.55	(1.43, 1.70)	1.77	(1.63, 1.93)	1.95	(1.76, 2.16)	2.41	(2.01, 2.87)
Lung	4.14	(3.64, 4.72)	4.75	(4.33, 5.20)	3.88	(3.56, 4.23)	4.69	(4.30, 5.12)	5.23	(4.51, 6.06)
Other/Multi	2.19	(1.87, 2.57)	1.78	(1.57, 2.01)	1.75	(1.56, 1.97)	2.09	(1.84, 2.38)	2.77	(2.25, 3.40)
Age; 1 year increase										
0-5	0.89	(0.82, 0.97)	0.94	(0.87, 1.01)	0.88	(0.81, 0.95)	0.86	(0.78, 0.95)	0.81	(0.68, 0.97)
5-19	1.12	(1.00, 1.24)	1.04	(0.95, 1.14)	1.14	(1.03, 1.26)	1.16	(1.03, 1.31)	1.25	(1.00, 1.57)
20+	1.03	(1.01, 1.06)	1.06	(1.04, 1.08)	1.03	(1.01, 1.06)	1.03	(1.00, 1.06)	1.02	(0.96, 1.08)

HR=hazard ratio, CI=confidence interval, NH=non-Hispanic; API=Asian/Pacific islander

Table 39: Relative risk of death by transplant year with all recipients censored at 3 years post-transplant

	HR	95% CI
Time since cancer		
<1 year	10.68	(10.36, 11.01)
1-3 years	3.41	(3.25, 3.58)
3-5 years	1.97	(1.82, 2.13)
5-7 years	1.66	(1.49, 1.85)
7-9 years	1.55	(1.34, 1.79)
9-15 years	1.43	(1.22, 1.68)
15+ years	1.14	(0.63, 2.07)
Sex		
Female vs. Male	0.95	(0.93, 0.96)
Race		
NH white	Ref.	
NH black	1.29	(1.26, 1.33)
Hispanic	0.93	(0.91, 0.96)
API	0.79	(0.75, 0.83)
Other	1.16	(1.04, 1.30)
Year of transplant		
1987-1994	Ref.	
1995-1999	0.87	(0.85, 0.90)
2000-2004	0.76	(0.74, 0.78)
2005-2008	0.67	(0.65, 0.69)
2009-2014	0.54	(0.52, 0.58)
Organ Transplanted		
Kidney	Ref.	
Liver	1.24	(1.21, 1.27)
Heart	1.37	(1.33, 1.41)
Lung	3.39	(3.28, 3.49)
Other/Multi	1.41	(1.36, 1.47)
Age; 1 year increase		
0-5	0.99	(0.96, 1.02)
6-19	1.01	(0.97, 1.05)
20+	1.04	(1.03, 1.05)

HR=hazard ratio, CI=confidence interval,
NH=non-Hispanic; API=Asian/Pacific islander

Table 40: Hazard ratios for associations with death among recipients with an interaction between cancer and time since cancer diagnosis

	HR	95% CI
Time since cancer		
<1 year	10.68	(10.36, 11.01)
1-3 years	3.41	(3.25, 3.58)
3-5 years	1.97	(1.82, 2.13)
5+ years	1.57	(1.45, 1.69)
Sex		
Female vs. Male	0.95	(0.93, 0.96)
Race		
NH White		
NH black	1.29	(1.26, 1.33)
Hispanic	0.93	(0.91, 0.96)
API	0.79	(0.75, 0.83)
Other	1.16	(1.04, 1.30)
Year of transplant		
1987-1994		
1995-1999	0.87	(0.85, 0.89)
2000-2004	0.76	(0.74, 0.78)
2005-2008	0.67	(0.65, 0.69)
2009-2014	0.54	(0.52, 0.58)
Organ transplanted		
Kidney		
Liver	1.24	(1.21, 1.27)
Heart	1.37	(1.33, 1.41)
Lung	3.39	(3.28, 3.49)
Other/Multi	1.41	(1.36, 1.47)
Age; 1 year increase		
0-5	0.99	(0.96, 1.02)
6-19	1.01	(0.97, 1.05)
20+	1.04	(1.03, 1.05)

-2LogL: 1084538.4
HR=hazard ratio, CI=confidence interval, NH=non-Hispanic; API=Asian/Pacific islander

Table 41: Hazard ratios for associations with death among recipients with an interaction between cancer and collapsed categories for time since cancer diagnosis

	p-value	HR	95% CI
Time since cancer			
<1 year	<.0001	12.89	(12.26, 13.55)
1-3 years	<.0001	3.74	(3.39, 4.11)
3-5 years	<.0001	2.16	(1.77, 2.63)
5+	<.0001	1.90	(1.48, 2.46)
Interaction between time since cancer and follow-up time			
<1 year	<.0001	0.97	(0.96, 0.97)
1-3 years	0.02	0.99	(0.97, 1.00)
3-5 years	0.24	0.99	(0.97, 1.01)
5+	0.07	0.98	(0.96, 1.00)
Sex			
Female vs. Male	<.0001	0.94	(0.93, 0.96)
Race			
NH White			
NH black	<.0001	1.29	(1.26, 1.33)
Hispanic	<.0001	0.93	(0.91, 0.96)
API	<.0001	0.79	(0.75, 0.83)
Other	0.01	1.16	(1.04, 1.30)
Year of transplant			
1987-1994			
1995-1999	<.0001	0.87	(0.85, 0.89)
2000-2004	<.0001	0.76	(0.74, 0.78)
2005-2008	<.0001	0.67	(0.65, 0.69)
2009-2014	<.0001	0.54	(0.52, 0.58)
Organ transplanted			
Kidney			
Liver	<.0001	1.24	(1.21, 1.27)
Heart	<.0001	1.37	(1.34, 1.41)
Lung	<.0001	3.38	(3.27, 3.49)
Other/Multi	<.0001	1.41	(1.36, 1.47)
Age at transplant			
0-5	0.35	0.99	(0.96, 1.02)
6-19	0.64	1.01	(0.97, 1.05)
20+	<.0001	1.04	(1.03, 1.05)

-2LogL: 1084451.5
HR=hazard ratio, SE=standard error, CI=confidence interval, NH=non-Hispanic; API=Asian/Pacific islander

Table 42: Relative risk of death among recipients with interaction between cancer effect and follow-up time

D. Sensitivity of the PAF to exclusions of cancers within one year after transplant

In the primary analysis, cancers were excluded if they occurred within 90 days of transplant. A sensitivity analysis was conducted to evaluate if using a one year exclusion window post-transplant impacted the results (Table 43). The PAF and cancer-attributable mortality rates are similar overall (516 vs. 518 per 100,000 person-years) and by subgroup. The largest absolute differences are in the time since transplant subgroups. This was expected since the largest differences between the two criteria occur in the time post-transplant. No other differences were found in the results.

	Results excluding cancers diagnosed within 90 days of transplant					Results excluding cancers diagnosed within 1 year of transplant				
	HR*	PAF (%)	95% CI	Overall mortality mate per 100,000 PY	Cancer-attributable mortality rate per 100,000 PY	HR*	PAF (%)	95% CI	Overall mortality rate per 100,000 PY	Cancer-attributable mortality Rate per 100,000 PY
Overall	4.79	13.2	(12.4, 13.9)	3922	516	4.74	13.8	(12.9, 14.6)	3767	518
Age										
0-17	5.58	6.1	(2.5, 9.8)	1740	107	9.08	6.8	(2.2, 11.4)	1514	103
18-34	5.56	7.3	(4.9, 9.7)	2055	151	5.9	7.4	(4.7, 10.1)	1987	147
35-49	5.09	10.5	(9.1, 11.9)	3134	329	5.1	10.9	(9.3, 12.4)	3069	333
50-64	4.7	15.5	(14.5, 16.5)	5277	818	4.58	16.4	(15.2, 17.6)	5085	833
65+	4.17	15.5	(13.8, 17.1)	7953	1229	4.11	16	(14.1, 17.9)	7789	1248
Sex										
Female	4.78	11.1	(10.0, 12.2)	3633	404	4.81	11.6	(10.3, 12.9)	3481	403
Male	4.79	14.4	(13.4, 15.3)	4113	590	4.69	15	(14.0, 16.1)	3957	594
Race										
Non-Hispanic white	4.78	14.2	(13.4, 15.1)	4212	600	4.73	14.9	(14.0, 15.9)	4040	602
Non-Hispanic black	4.11	10.2	(8.5, 11.8)	4171	423	4.07	10.5	(8.6, 12.5)	4065	429
Hispanic	5.61	10.7	(8.3, 13.0)	2850	303	5.69	11.1	(8.3, 13.9)	2713	301
Asian/Pacific islander	7.34	13.8	(9.0, 18.6)	2533	350	6.69	13.9	(8.2, 19.6)	2351	327
Other	6.13	12.4	(3.4, 21.4)	4080	506	5.39	12.6	(2.4, 22.8)	3919	495
Organ										
Kidney	4.67	13.1	(11.9, 14.3)	3141	410	4.56	13.2	(11.9, 14.5)	3154	415
Liver	5.63	13	(11.6, 14.5)	4080	531	5.58	14.2	(12.4, 16.0)	3657	519
Heart	4.6	16.4	(14.8, 17.9)	5439	891	4.63	17.6	(15.9, 19.4)	5174	912
Lung	4.04	10.5	(9.2, 11.8)	11836	1241	4.35	11.2	(9.6, 12.7)	11290	1259
Other	5.8	9.8	(7.0, 12.6)	3491	342	6.05	10.5	(7.1, 13.8)	3151	330

Table 43: Population-attributable fraction and mortality rates after excluding cancers 90 days and 1 year post-transplant

Table 43: Continued

	Results excluding cancers diagnosed within 90 days of transplant					Results excluding cancers diagnosed within 1 year of transplant				
	HR*	PAF (%)	95% CI	Overall mortality mate per 100,000 PY	Cancer-attributable mortality rate per 100,000 PY	HR*	PAF (%)	95% CI	Overall mortality rate per 100,000 PY	Cancer-attributable mortality Rate per 100,000 PY
Years since transplant										
0-1.99	11.6	8.1	(6.4, 9.9)	3853	313	11.54	8.7	(6.3, 11.0)	3105	269
2-4.99	6.27	13.5	(11.3, 15.7)	3371	456	6.47	14.3	(12.0, 16.6)	3515	503
5-9.99	4.04	16	(14.1, 18.0)	4199	673	3.88	15.7	(13.7, 17.8)	4371	688
10+	2.61	15.7	(13.0, 18.4)	5154	810	2.61	16.1	(12.9, 19.2)	5217	838
Cancer Site										
Lung	10.42	3.1	(2.8, 3.4)	3922	121	10.23	3.4	(3.0, 3.8)	3767	127
NHL	4.58	1.9	(1.6, 2.2)	3922	75	4.73	1.7	(1.4, 2.0)	3767	65
CRC	3.48	0.7	(0.6, 0.9)	3922	29	3.64	0.8	(0.6, 1.0)	3767	31
Breast (men and women)	1.79	0.3	(0.1, 0.4)	3922	10	1.64	0.2	(0.1, 0.4)	3767	9
Kidney	2.42	0.5	(0.3, 0.6)	3922	18	2.35	0.5	(0.3, 0.6)	3767	18
Melanoma	2.26	0.3	(0.2, 0.4)	3922	11	2.14	0.3	(0.2, 0.4)	3767	11
Liver	13.35	0.3	(0.2, 0.4)	3922	10	13.44	0.3	(0.2, 0.4)	3767	11
Thyroid	1.4	0	(0.0, 0.1)	3922	2	1.19	0	(0.0, 0.1)	3767	1
Prostate	0.99	-0.01	--	3922	-1	0.97	-0.04	--	3767	-1
Infectious etiology†	4.81	3	(2.6, 3.3)	3922	116	5.05	2.8	(2.5, 3.2)	3767	107
Non-Infectious etiology	4.3	9.9	(9.3, 10.5)	3922	388	4.22	10.6	(9.9, 11.3)	3767	399

p_d=Proportion of all deaths among those with cancer; HR=Hazard ratio; PAF=Population-attributable fraction; CI=Confidence interval; PY=Person-

*Models adjusted for: sex, race, organ transplanted, age at transplant and year of transplant

†NHL, liver, stomach, Kaposi sarcoma, oropharynx including tonsil, anus, Hodgkin lymphoma, vulva, cervix, penis, nasopharynx, vagina

Appendix 2: Statistical considerations in estimating the life-years lost to cancer

A. Selection of censoring at 10 years post-transplant

The distribution of follow-up time is shown in Figure 29 and the number of deaths in each time interval post-transplant are shown in Table 44. Although, there are deaths occurring at years 15 and beyond, I encountered problems with the estimation for various subgroups. Full follow-up data were used for the estimation of the Cox regression models in aim 2.1 and 2.2. However, estimation of the LYL to cancer was censored at 10 years to avoid extrapolation beyond the data into the tail.

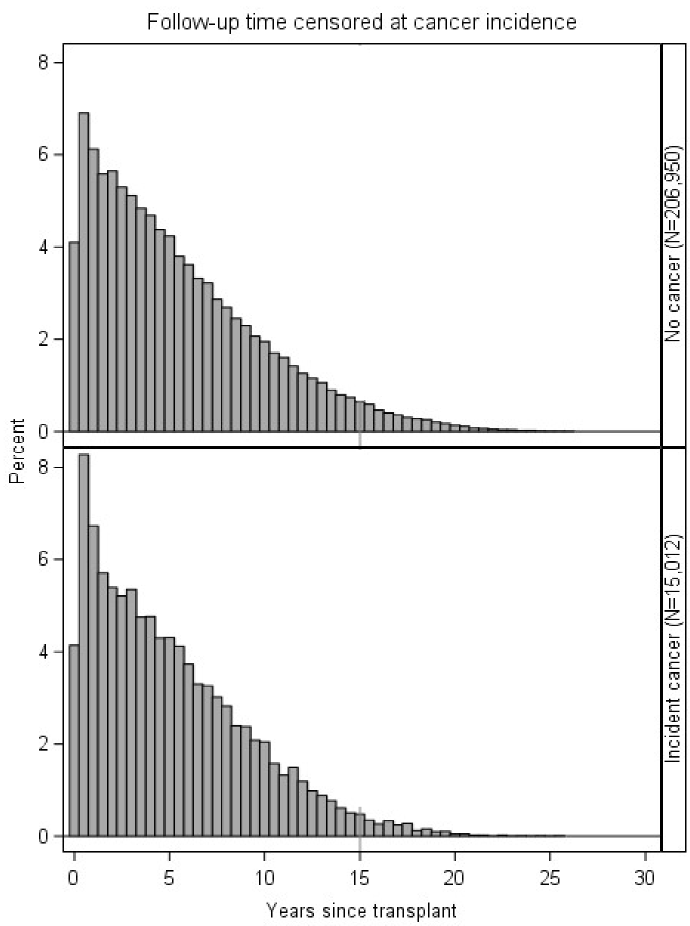


Figure 29: Distribution of follow-up time among recipients who developed incident cancer and those who did not

Follow-up time, years	**0-4.99**	**5-9.99**	**10-14.99**	**15-19.99**	**20-24.99**	**25-29.99**
All organs	24892	10943	3953	861	59	0
Kidney	11308	5735	1826	310	19	0
Liver	5192	2043	890	247	18	0
Heart	3623	1764	813	226	12	0
Lung	3296	843	195	25	1	0
Other/Multi	1473	558	229	53	9	0

Table 44: Deaths at each time interval with censoring at cancer incidence

B. Functional form of continuous covariates

There were two covariates that could be modeled as continuous: Age at transplant and time from initial dialysis to transplant for kidney recipients.

Age at transplant

Age is modeled with a piecewise linear spline as described in Appendix 1, section A.

Time from initial dialysis to transplant

Years from initial dialysis to transplant was categorized into 8 categories and the estimated beta coefficients were plotted (Figure 30). Time from initial dialysis was modeled as categorical to accommodate the missing and No dialysis categories.

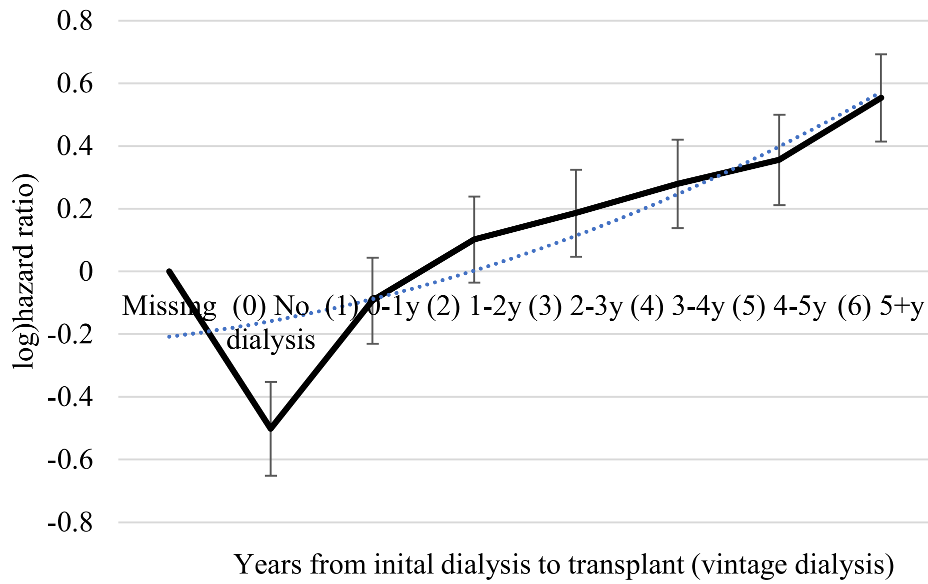


Figure 30: Estimates log hazard ratios for years from initial dialysis to transplant among kidney recipients

C. Selection of caliper width for matching

The caliper used for matching was the distance in the prognostic score between a cancer case and a potential control as described in Chapter 7. The number of matches available for several caliper widths were examined (Table 45). Note that these case counts are before final exclusions performed in the final analysis and shown in Table 24. A difference of 0.2 was selected as the optimal width since there appears to be a drop in the number of cases without available controls compared to 0.01 or 0.1. Yet, this distance seems reasonable as it represents a 22% difference in the risk between cases and matched controls (i.e. $e^{0.2} = 1.22$).

				Controls per case						
Organ transplanted	**Difference in linear predictor**	**Number of cancer cases**	**Number of cases**	**1**	**2**	**3**	**4**	**5**	**5+**	**No controls**
Kidney	0.01	8028		971	860	797	718	611	2661	1098
	0.1			109	118	124	127	121	728	134
	0.18			54	65	65	72	71	407	72
	0.25			32	51	42	44	60	315	56
	0.3			30	29	50	40	39	257	44
	0.4			31	19	34	29	24	185	29
	0.5			18	18	17	22	29	151	24
	Any distance				8	9	10	12	8	68
Range of difference: 0-3.29										
Liver	0.01	2781		421	309	233	198	168	628	650
	0.1			85	72	70	83	77	431	84
	0.2			41	52	42	47	54	252	33
	0.3			16	29	32	35	29	183	22
	0.4			10	15	25	25	20	126	16
	0.5			10	11	19	20	17	104	13
	Any distance				8	4	7	9	14	67
Range of difference: 0-2.07										
Heart	0.01	2628		338	240	211	177	136	562	445
	0.1			79	92	84	70	77	394	93
	0.2			45	42	53	53	36	235	46
	0.3			33	24	35	25	39	193	25
	0.4			26	20	26	17	30	158	19
	0.5			20	17	23	16	24	128	16
	Any distance				14	9	16	13	20	100
Range of difference: 0-1.7										
Lung	0.01	942		180	96	67	63	58	164	333
	0.1			44	46	29	35	35	200	55
	0.2			26	27	24	21	28	144	27
	0.3			24	22	14	21	22	119	20
	0.4			20	19	12	20	20	115	19
	0.5			18	21	12	17	19	113	19
	Any distance				18	21	12	17	18	111
Range of difference: 0-0.83										

Table 45: Number of controls per case for a range of caliper widths

Table 45 continued

Organ transplanted	Difference in linear predictor	Number of cancer cases	Number of cases	Controls per case 1	2	3	4	5	5+
Other/Multi									
0.01	633		80	49	36	39	26	116	161
0.1			44	34	36	19	25	104	63
0.2			32	20	29	23	13	98	29
0.3			26	20	21	23	10	93	19
0.4			24	14	22	20	8	83	13
0.5			17	16	21	19	8	82	13
Any distance				12	13	13	18	5	78

Range of difference: 0-1.12

References

1. Engels EA, Pfeiffer RM, Fraumeni JF, Jr., et al: Spectrum of cancer risk among US solid organ transplant recipients. JAMA 306:1891-901, 2011

2. Acuna SA, Fernandes KA, Daly C, et al: Cancer Mortality Among Recipients of Solid-Organ Transplantation in Ontario, Canada. JAMA Oncol 2:463-9, 2016

3. Cheung CY, Chu KH, Chow KM, et al: Malignancies after kidney transplantation: Hong Kong renal registry. Am J Transplant 12:3039-46, 2012

4. Farrugia D, Mahboob S, Cheshire J, et al: Malignancy-related mortality following kidney transplantation is common. Kidney Int 85:1395-403, 2014

5. Na R, Grulich AE, Meagher NS, et al: De novo cancer-related death in Australian liver and cardiothoracic transplant recipients. Am J Transplant 13:1296-1304, 2013

6. US Department of Health and Human Services, United Network for Organ Sharing,: Organ Procurement and Transplantation Network,

7. US Department of Health and Human Services, United Network for Organ Sharing,: How Organ Allocation Works. https://optn.transplant.hrsa.gov/learn/about-transplantation/how-organ-allocation-works/

8. Valapour M, Skeans MA, Smith JM, et al: OPTN/SRTR 2015 Annual Data Report: Lung. Am J Transplant 17 Suppl 1:357-424, 2017

9. Kim WR, Lake JR, Smith JM, et al: OPTN/SRTR 2015 Annual Data Report: Liver. American Journal of Transplantation 17:174-251, 2017

10. Hart A, Smith JM, Skeans MA, et al: OPTN/SRTR 2015 Annual Data Report: Kidney. Am J Transplant 17 Suppl 1:21-116, 2017

11. Colvin M, Smith JM, Skeans MA, et al: OPTN/SRTR 2015 Annual Data Report: Heart. Am J Transplant 17 Suppl 1:286-356, 2017

12. Department of Health and Human Services, Health Resources and Services Administration,, Healthcare Systems Bureau,, Division of Transplantation: OPTN/SRTR 2003 Annual Report of the U.S. Organ procurement and Transplantation Network and the Scientific Registry of Transplant Recepients.

Rockville, MD; United Network for Organ Sharing, Richmond, VA; University Renal Research and Education Association, Ann Arbor, MI., 2003

13. Abadie A, Imbens GW: On the failure of the bootsrap for matching estimators. Econometrica 76:1537-1557, 2008

14. National, Cancer,,Institute: Risk Factors,

15. Safaeian M, Robbins HA, Berndt SI, et al: Risk of Colorectal Cancer After Solid Organ Transplantation in the United States. Am J Transplant 16:960-7, 2016

16. Karami S, Yanik EL, Moore LE, et al: Risk of Renal Cell Carcinoma Among Kidney Transplant Recipients in the United States. Am J Transplant 16:3479-3489, 2016

17. Minai OA, Shah S, Mazzone P, et al: Bronchogenic carcinoma after lung transplantation: characteristics and outcomes. J Thorac Oncol 3:1404-9, 2008

18. Gutierrez-Dalmau A, Campistol JM: Immunosuppressive therapy and malignancy in organ transplant recipients: a systematic review. Drugs 67:1167-98, 2007

19. Dunn GP, Bruce AT, Ikeda H, et al: Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol 3:991-998, 2002

20. Shinkai Y, Rathbun G, Lam KP, et al: RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell 68:855-67, 1992

21. Herrera LA, Benítez-Bribiesca L, Mohar A, et al: Role of infectious diseases in human carcinogenesis. Environmental and Molecular Mutagenesis 45:284-303, 2005

22. Kuper H, Adami HO, Trichopoulos D: Infections as a major preventable cause of human cancer. Journal of Internal Medicine 248:171-183, 2000

23. Koshiol J, Pawlish K, Goodman MT, et al: Risk of Hepatobiliary Cancer After Solid Organ Transplant in the United States. Clinical Gastroenterology and Hepatology 12:1541-1549.e3, 2014

24. Kiberd BA, Rose C, Gill JS: Cancer mortality in kidney transplantation. Am J Transplant 9:1868-75, 2009

25. Morgenstern HAL, Kleinbaum DG, Kupper LL: Measures of Disease Incidence Used in Epidemiologic Research. International Journal of Epidemiology 9:97-104, 1980

26. Gardner JW, Sanborn JS: Years of potential life lost (YPLL)--what does it measure? Epidemiology 1:322-9, 1990

27. Begg CB, Schrag D: Attribution of deaths following cancer treatment. J Natl Cancer Inst 94:1044-5, 2002

28. Rockhill B, Newman B, Weinberg C: Use and misuse of population attributable fractions. Am J Public Health 88:15-9, 1998

29. Simard EP, Pfeiffer RM, Engels EA: Mortality due to cancer among people with AIDS: a novel approach using registry-linkage data and population attributable risk methods. AIDS 26:1311-8, 2012

30. Engels EA, Yanik EL, Wheeler W, et al: Cancer-Attributable Mortality Among People With Treated Human Immunodeficiency Virus Infection in North America. Clinical Infectious Diseases 65:636-643, 2017

31. Horner MJ, Shiels MS, Pfeiffer RM, et al: Deaths attributable to cancer in the United States HIV population during 2001-2015. Clin Infect Dis, 2020

32. Percy C, Stanek E, 3rd, Gloeckler L: Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Public Health 71:242-50, 1981

33. German RR, Fink AK, Heron M, et al: The accuracy of cancer mortality statistics based on death certificates in the United States. Cancer Epidemiology 35:126-131, 2011

34. Welch HG, Black WC: Are deaths within 1 month of cancer-directed surgery attributed to cancer? J Natl Cancer Inst 94:1066-70, 2002

35. Brown BW, Brauner C, Levy LB: Assessing changes in the impact of cancer on population survival without considering cause of death.

36. Chu KC, Miller BA, Feuer EJ, et al: A method for partitioning cancer mortality trends by factors associated with diagnosis: an application to female breast cancer. J Clin Epidemiol 47:1451-61, 1994

37. Howlader N, Ries LA, Mariotto AB, et al: Improved estimates of cancer-specific survival rates from population-based data. J Natl Cancer Inst 102:1584-98, 2010

38. Wolfe RA, McCullough KP, Schaubel DE, et al: Calculating life years from transplant (LYFT): methods for kidney and kidney-pancreas candidates. Am J Transplant 8:997-1011, 2008

39. Arbogast PG, Ray WA: Performance of disease risk scores, propensity scores, and traditional multivariable outcome regression in the presence of multiple confounders. Am J Epidemiol 174:613-20, 2011

40. Hansen BB: The prognostic analogue of the propensity score. Biometrika 95:481-488, 2008

41. Miettinen OS: Stratification by a multivariate confounder score. Am J Epidemiol 104:609-20, 1976

42. Li Y, Schaubel DE, He K: Matching methods for obtaining survival functions to estimate the effect of a time-dependent treatment. Stat Biosci 6:105-126, 2014

43. Zucker DM: Restricted Mean Life with Covariates: Modification and Extension of a Useful Survival Analysis Method. Journal of the American Statistical Association 93:702-709, 1998

44. Chen PY, Tsiatis AA: Causal inference on the difference of the restricted mean lifetime between two groups. Biometrics 57:1030-8, 2001

45. Zhang M, Schaubel DE: Estimating differences in restricted mean lifetime using observational data subject to dependent censoring. Biometrics 67:740-9, 2011

46. Doll R, Boreham J: Recent trends in cancer mortality in the UK. Br J Cancer 92:1329-35, 2005

47. Lortet-Tieulent J, Soerjomataram I, Lin CC, et al: U.S. Burden of Cancer by Race and Ethnicity According to Disability-Adjusted Life Years. Am J Prev Med 51:673-681, 2016

48. Acuna SA, Huang JW, Scott AL, et al: Cancer Screening Recommendations for Solid Organ Transplant Recipients: A Systematic Review of Clinical Practice Guidelines. American Journal of Transplantation 17:103-114, 2016

49. National, Cancer,,Institute: SEER Statistics, 2017

50. Yanik EL, Nogueira LM, Koch L, et al: Comparison of cancer diagnoses between the US solid organ transplant registry and linked central cancer registries. Am J Transplant, 2016

51. Surveillance, Epidemiology and End Results Program; National Cancer Institute,: Incidence Site Recode Variables,

52. SCHOENFELD D: Chi-squared goodness-of-fit tests for the proportional hazards regression model. Biometrika 67:145-153, 1980

53. Little RJA, Rubin DB: Statistical Analysis with Missing Data. United Kingdom, Wiley, 2019

54. Jewell NP: Statistics for Epidemiology. London, Chapman & Hall/CRC, 2004 pp. 84-85

55. Mark Stevenson with contributions from Telmo Nunes, Cord Heuer,, Jonathon Marshall,, Javier Sanchez,, Ron Thornton,, Jeno Reiczigel,, Jim Robison-Cox,, Paola Sebastiani,, Peter Solymos,, Kazuki Yoshida,, Geoff Jones,, Sarah Pirikahu,, Simon Firestone and Ryan Kyle,: epiR: Tools for the Analysis of Epidemiological Data. R package version 0.9-79, 2016

56. Garrett GL, Blanc PD, Boscardin J, et al: Incidence of and risk factors for skin cancer in organ transplant recipients in the united states. JAMA Dermatology 153:296-303, 2017

57. Ulrich C, Kanitakis J, Stockfleth E, et al: Skin Cancer in Organ Transplant Recipients—Where Do We Stand Today? American Journal of Transplantation 8:2192-2198, 2008

58. Connolly JG, Gagne JJ: Comparison of Calipers for Matching on the Disease Risk Score. Am J Epidemiol 183:937-48, 2016

59. Wyss R, Ellis AR, Brookhart MA, et al: Matching on the disease risk score in comparative effectiveness research of new treatments. Pharmacoepidemiol Drug Saf 24:951-61, 2015

60. Austin PC: Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm Stat 10:150-61, 2011

61. Rosenbaum PR, Rubin DB: Constructing a control group using multivariate matched sampling methods that incorportate the propensity score. The American Statistician 39:33-38, 1985

62. D'Arcy ME, Coghill AE, Lynch CF, et al: Survival after a cancer diagnosis among solid organ transplant recipients in the United States. Cancer, 2019

63. Miao Y, Everly JJ, Gross TG, et al: De novo cancers arising in organ transplant recipients are associated with adverse outcomes compared with the general population. Transplantation 87:1347-59, 2009

64. Graubard BI, Fears TR: Standard errors for attributable risk for simple and complex sample designs. Biometrics 61:847-55, 2005

65. Cohen J: A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20:37-46, 1960

66. Na R, Grulich AE, Meagher NS, et al: Comparison of de novo cancer incidence in Australian liver, heart and lung transplant recipients. Am J Transplant 13:174-83, 2013

67. Collett D, Mumford L, Banner NR, et al: Comparison of the incidence of malignancy in recipients of different types of organ: a UK Registry audit. Am J Transplant 10:1889-96, 2010

68. Allen UD, Preiksaitis JK, Practice ASTIDCo: Epstein-Barr virus and posttransplant lymphoproliferative disorder in solid organ transplantation. Am J Transplant 13 Suppl 4:107-20, 2013

69. Yanik EL, Shiels MS, Smith JM, et al: Contribution of solid organ transplant recipients to the pediatric non-hodgkin lymphoma burden in the United States. Cancer 123:4663-4671, 2017

70. Yanik EL, Smith JM, Shiels MS, et al: Cancer Risk After Pediatric Solid Organ Transplantation. Pediatrics 139, 2017

71. VanderWeele TJ, Ding P: Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med 167:268-274, 2017

72. Negoita S, Feuer EJ, Mariotto A, et al: Annual Report to the Nation on the Status of Cancer, part II: Recent changes in prostate cancer trends and disease characteristics. Cancer 124:2801-2814, 2018

73. Shiels MS, Goedert JJ, Moore RD, et al: Reduced risk of prostate cancer in U.S. Men with AIDS. Cancer Epidemiol Biomarkers Prev 19:2910-5, 2010

74. Fellner MD, Durand KA, Solernou V, et al: Epstein-Barr virus load in transplant patients: Early detection of post-transplant lymphoproliferative disorders. Rev Argent Microbiol 48:110-8, 2016

75. Suraweera D, Sundaram V, Saab S: Treatment of Hepatitis C Virus Infection in Liver Transplant Recipients. Gastroenterol Hepatol (N Y) 12:23-30, 2016

76. Terhorst D, Drecoll U, Stockfleth E, et al: Organ transplant recipients and skin cancer: assessment of risk factors with focus on sun exposure. Br J Dermatol 161 Suppl 3:85-9, 2009

77. Steenland K, Greenland S: Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. Am J Epidemiol 160:384-92, 2004

78. Hurst FP, Altieri M, Patel PP, et al: Effect of smoking on kidney transplant outcomes: analysis of the United States Renal Data System. Transplantation 92:1101-7, 2011

79. Leithead JA, Ferguson JW, Hayes PC: Smoking-related morbidity and mortality following liver transplantation. Liver Transpl 14:1159-64, 2008

80. Weinrauch LA, Claggett B, Liu J, et al: Smoking and outcomes in kidney transplant recipients: a post hoc survival analysis of the FAVORIT trial. Int J Nephrol Renovasc Dis 11:155-164, 2018

81. Grulich AE, van Leeuwen MT, Falster MO, et al: Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta-analysis. Lancet 370:59-67, 2007

82. Hall EC, Pfeiffer RM, Segev DL, et al: Cumulative incidence of cancer after solid organ transplantation. Cancer 119:2300-8, 2013

83. Noone AM, Pfeiffer RM, Dorgan JF, et al: Cancer-attributable mortality among solid organ transplant recipients in the United States: 1987 through 2014. Cancer 125:2647-2655, 2019

84. Makuch RW: Adjusted Survival Curve Estimation Using Covariates. Journal of Chronic Diseases 35:437–443, 1982

85. Qi T, Wang J: Calcuating Restricted Mean Survival Time, in PharmaSUG (ed), 2018

86. Ragunathan TE, Solenberger PW, Van Hoewyk J: IVEware: Imputatin and Variance Estimation Software. Survey Methodology Program, Survey Research Center, Institute for Social Research, University of Michigan, 2002

87. Triplette M, Crothers K, Mahale P, et al: Risk of lung cancer in lung transplant recipients in the United States. Am J Transplant 19:1478-1490, 2019

88. Rosales BM, De La Mata N, Vajdic CM, et al: Cancer mortality in kidney transplant recipients: An Australian and New Zealand population-based cohort study, 1980-2013. Int J Cancer 146:2703-2711, 2020

89. Moyer VA, Force USPST: Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 160:330-8, 2014

90. Anis KH, Weinrauch LA, D'Elia JA: Effects of Smoking on Solid Organ Transplantation Outcomes. Am J Med 132:413-419, 2019

91. Mangus RS, Fridell JA, Kubal CA, et al: Worse Long-term Patient Survival and Higher Cancer Rates in Liver Transplant Recipients With a History of Smoking. Transplantation 99:1862-8, 2015

92. Corbett C, Armstrong MJ, Neuberger J: Tobacco Smoking and Solid Organ Transplantation. Transplantation 94:979-987, 2012

93. Takahashi A, Hu SL, Bostom A: Physical Activity in Kidney Transplant Recipients: A Review. Am J Kidney Dis 72:433-443, 2018

94. Kallwitz ER, Loy V, Mettu P, et al: Physical activity and metabolic syndrome in liver transplant recipients. Liver Transpl 19:1125-31, 2013

95. Myers J, Gullestad L, Bellin D, et al: Physical activity patterns and exercise performance in cardiac transplant recipients. J Cardiopulm Rehabil 23:100-6, 2003

96. Pang A, Lingham S, Zhao W, et al: Physician Practice Patterns and Barriers to Counselling on Physical Activity in Solid Organ Transplant Recipients. Ann Transplant 23:345-359, 2018

97. Eheman C, Henley SJ, Ballard-Barbash R, et al: Annual Report to the Nation on the status of cancer, 1975-2008, featuring cancers associated with excess weight and lack of sufficient physical activity. Cancer 118:2338-2366, 2012

98. Lee BP, Terrault NA: Return to Alcohol Use After Liver Transplant: Patterns and Surveillance. Clinical Liver Disease 12:160-164, 2018

99. Henley SJ, Thomas CC, Lewis DR, et al: Annual report to the nation on the status of cancer, part II: Progress toward Healthy People 2020 objectives for 4 common cancers. Cancer 126:2250-2266, 2020

100. Leveridge M, Musquera M, Evans A, et al: Renal cell carcinoma in the native and allograft kidneys of renal transplant recipients. J Urol 186:219-23, 2011

101. Tsaur I, Obermüller N, Jonas D, et al: De novo renal cell carcinoma of native and graft kidneys in renal transplant recipients. BJU Int 108:229-34, 2011

102. Neuzillet Y, Lay F, Luccioni A, et al: De novo renal cell carcinoma of native kidney in renal transplant recipients. Cancer 103:251-7, 2005

103. Dickson RP, Davis RD, Rea JB, et al: High frequency of bronchogenic carcinoma after single-lung transplantation. J Heart Lung Transplant 25:1297-301, 2006

104. Mathew J, Kratzke RA: Lung Cancer and Lung Transplantation: A Review. Journal of Thoracic Oncology 4:753-760, 2009

105. Barrett WL, First MR, Aron BS, et al: Clinical course of malignancies in renal transplant recipients. Cancer 72:2186-9, 1993

106. Robbins HA, Clarke CA, Arron ST, et al: Melanoma Risk and Survival among Organ Transplant Recipients. J Invest Dermatol 135:2657-2665, 2015

107. Ge F, Li C, Xu X, et al: Cancer risk in heart or lung transplant recipients: A comprehensive analysis of 21 prospective cohorts. Cancer Med 9:9595-9610, 2020

108. Vajdic CM, McDonald SP, McCredie MR, et al: Cancer incidence before and after kidney transplantation. Jama 296:2823-31, 2006

109. Villeneuve PJ, Schaubel DE, Fenton SS, et al: Cancer incidence among Canadian kidney transplant recipients. Am J Transplant 7:941-8, 2007

110. Jiang Y, Villeneuve PJ, Wielgosz A, et al: The incidence of cancer in a population-based cohort of Canadian heart transplant recipients. Am J Transplant 10:637-45, 2010

111. Vajdic CM, van Leeuwen MT: Cancer incidence and risk factors after solid organ transplantation. Int J Cancer 125:1747-54, 2009

112. Jiang Y, Villeneuve PJ, Fenton SS, et al: Liver transplantation and subsequent risk of cancer: findings from a Canadian cohort study. Liver Transpl 14:1588-97, 2008

113. Howlader N, Noone AM, Krapcho M, et al: SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2013/, based on November 2015 SEER data submission, posted to the SEER web site, April 2016.,

114. Allen UD, Preiksaitis JK, Practice tAIDCo: Post-transplant lymphoproliferative disorders, Epstein-Barr virus infection, and disease in solid organ transplantation: Guidelines from the American Society of Transplantation Infectious Diseases Community of Practice. Clinical Transplantation 33:e13652, 2019

115. Green M, Michaels MG: Epstein–Barr Virus Infection and Posttransplant Lymphoproliferative Disorder. American Journal of Transplantation 13:41-54, 2013

116. Werner JM, Vermehren J: New HCV Therapies and Liver Transplantation. Transplantation 100:260-261, 2016

117. Hartman RI, Green AC, Gordon LG, et al: Sun Protection Among Organ Transplant Recipients After Participation in a Skin Cancer Research Study. JAMA Dermatology 154:842-844, 2018

118. Wu SZ, Jiang P, DeCaro JE, et al: A qualitative systematic review of the efficacy of sun protection education in organ transplant recipients. J Am Acad Dermatol 75:1238-1244.e5, 2016

119. Harrell F: SAS Macros for Assisting with Survival and Risk Analysis, and Some SAS Procedures Useful for Multivariable Modeling. Vanderbilt Biostatistics Wiki, http://biostat.mc.vanderbilt.edu/wiki/Main/SasMacros, 2004

120. Stone CJ, Koo CY: Additive splines in statistics. Proc Stat Comp Sect Am Statist Assoc:45-8, 1985

121. Gassama M, Benichou J, Dartois L, et al: Comparison of methods for estimating the attributable risk in the context of survival analysis. BMC Med Res Methodol 17:10, 2017

122. Spiegelman D, Hertzmark E, Wand HC: Point and interval estimates of partial population attributable risks in cohort studies: examples and software. Cancer Causes Control 18:571-9, 2007

123. Benichou J: A review of adjusted estimators of attributable risk. Stat Methods Med Res 10:195-216, 2001