



Published in final edited form as:

*Endocrine*. 2021 December ; 74(3): 592–602. doi:10.1007/s12020-021-02779-x.

## Cardiovascular outcomes and rates of fractures and falls among patients with brand-name versus generic L-thyroxine use

Juan P. Brito<sup>1</sup>,  
Joseph S. Ross<sup>2,3,4</sup>,  
Yihong Deng<sup>5</sup>,  
Lindsey Sangaralingham<sup>5</sup>,  
David J. Graham<sup>6</sup>,  
Yandong Qiang<sup>6</sup>,  
Zhong Wang<sup>7</sup>,  
Xiaoxi Yao<sup>5,8</sup>,  
Liang Zhao<sup>7</sup>,  
Robert C. Smallridge<sup>9</sup>,  
Victor Bernet<sup>9</sup>,  
Nilay D. Shah<sup>5,8,10</sup>,  
Kasia J. Lipska<sup>4,11</sup>

<sup>1</sup>Division of Endocrinology, Diabetes, Metabolism, Nutrition, Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA

<sup>2</sup>Section of General Internal Medicine and the National Clinician Scholars Program, Yale School of Medicine, New Haven, CT, USA

<sup>3</sup>Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA

<sup>4</sup>Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, CT, USA

✉ Juan P. Brito, brito.juan@mayo.edu.

**Author contributions** J.P.B., J.S.R., L.S., N.D.S., and K.J.L. contributed to the study concept and design. J.P.B., J.S.R., Y.D., L.S., D.J. G., Y.Q., Z.W., X.Y., L.Z., N.D.S., and K.J.L. contributed to the development of the study protocol. Y.D. and L.S. were responsible for conducting the data analysis for this study. All authors contributed to the data interpretation. All authors critically reviewed the manuscript.

Compliance with ethical standards

**Conflict of interest** J.S.R. has received research support through Yale University from Johnson and Johnson to develop methods of clinical trial data sharing, from the Medical Device Innovation Consortium as part of the National Evaluation System for Health Technology (NEST), from the Agency for Healthcare Research and Quality (R01HS022882), from the National Heart, Lung and Blood Institute of the National Institutes of Health (NIH) (R01HS025164), and from the Laura and John Arnold Foundation to establish the Good Pharma Scorecard at Bioethics International and to establish the Collaboration for Research Integrity and Transparency (CRIT) at Yale. N.D.S. has received research support through Mayo Clinic from the Centers of Medicare and Medicaid Innovation, from the Agency for Healthcare Research and Quality (R01HS025164; R01HS025402; R03HS025517; U19HS024075), from the National Heart, Lung and Blood Institute of the National Institutes of Health (NIH) (R56HL130496; R01HL131535), National Science Foundation, and from the Patient Centered Outcomes Research Institute (PCORI). K.J.L. receives support from the Centers of Medicare and Medicaid Services (CMS) and the National Institute on Aging and the American Federation of Aging Research through the Paul Beeson Career Development Award (K23AG048359). R.C.S. receives support from the Alfred D. and Audrey M. Petersen Professorship in Cancer Research. D. J.G., Y.Q., Z.W., and L.Z. are employed by the US Food and Drug Administration. The remaining authors have nothing to disclose.

<sup>5</sup>Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, USA

<sup>6</sup>Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, USA

<sup>7</sup>Office of Research and Standards, Office of Generic Drugs, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, USA

<sup>8</sup>Division of Health Care Policy & Research, Mayo Clinic, Rochester, MN, USA

<sup>9</sup>Division of Endocrinology, Mayo Clinic, Jacksonville, FL, USA

<sup>10</sup>Optum Labs, Cambridge, MA, USA

<sup>11</sup>Section of Endocrinology, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA

## Abstract

**Purpose**—To compare cardiovascular outcomes and rates of fractures and falls among patients with persistent brand-name versus generic L-thyroxine use.

**Methods**—Retrospective, 1:1 propensity-matched longitudinal study using a national administrative claims database to examine adults (> 18 years) who initiated either brand or generic L-thyroxine between 2008 and 2018, censored at switch or discontinuation of L-thyroxine formulation or disenrollment from the health plan. Main outcome measures included rates of hospitalization for atrial fibrillation, myocardial infarction, congestive heart failure, stroke, spine and hip fractures, and rate of falls in the outpatient or inpatient setting. Hospitalizations for pneumonia were used as a negative control.

**Results**—195,046 adults initiated treatment with L-thyroxine between 2008 and 2017: 87% generic and 13% brand formulations. They were mostly women (76%), young (94.6% under age 65), white (66%), and 47% had baseline thyroid stimulating hormone levels between 4.5 and 9.9 mIU/L. Among 35,667 propensity-matched patients, there were no significant differences between patients treated with brand versus generic L-thyroxine in atrial fibrillation (HR 0.96, 0.58–1.60), myocardial infarction (HR 0.66, 0.39–1.14), congestive heart failure (HR 1.30, 0.78–2.16), stroke (0.72, 0.49–1.06), spine (HR 0.87, 0.38–1.99) and hip fractures (HR 0.86, 0.26–2.82), or fall outcomes (HR 1.02, 0.14–7.32). Hospitalization rates for pneumonia (used as negative control) did not differ between groups (HR 0.85, 0.61–1.19). There were no interactions between brand versus generic L-thyroxine, these outcomes, and thyroid cancer, age, or L-thyroxine dose subgroups.

**Conclusions**—We found no significant differences in cardiovascular outcomes and rates of falls and fractures for patients who filled brand versus generic L-thyroxine.

## Keywords

Generic; Brand; Levothyroxine; Hypothyroidism

## Background

L-thyroxine is the most common thyroid hormone replacement therapy used for patients with hypothyroidism and it has become the most frequently prescribed medication in the United States [1, 2]. L-thyroxine is a synthetic form of thyroxine and is available as a brand-name or generic preparation. Generic formulations are chemically identical to their reference brand-name drugs in dosage form, strength, route of administration, quality, performance characteristics, and intended use, but they are substantially cheaper than brand-name L-thyroxine [3].

Despite potential advantages of prescribing generic L-thyroxine, the use of brand L-thyroxine continues to be common. Approximately 40% of L-thyroxine formulations prescribed by endocrinologists in 2016 were for brand L-thyroxine [4]. Once patients received a brand L-thyroxine prescription, they are more likely to continue taking that formulation. One study found that only 20% of brand L-thyroxine initiators experienced generic substitution within 12 months [5].

L-thyroxine is a drug with a narrow therapeutic index, meaning that small deviation of the dose and subsequent drug blood concentration can cause clinically meaningful effects on patients including therapeutic failures or adverse reactions. Some experts have warned that small deviation of drug blood concentrations could arise from using generic L-thyroxine formulations as opposed to consistent use of brand L-thyroxine [6, 7]. A possible difference in clinical effect across L-thyroxine formulations could cause abnormal levels of thyroid hormone concentration in the blood that may affect metabolic demands of the heart and bone, and possibly result in heart failure exacerbation, heart ischemia, arrhythmia, and bone loss [8, 9]. A study by Smallridge et al., using national data from a large administrative claims database, found no difference in the incidence of cardiovascular events between patients using generic versus brand L-thyroxine. This study, however, did not adjust for baseline thyroid hormone values, the dose of L-thyroxine, or assess rates of falls and fractures [10].

To better understand the comparative safety of branded and generic L-thyroxine, we used a national administrative database to explore the impact of persistent use on cardiovascular events and on rates of falls and fractures, after propensity-matching the two groups. Unlike the prior study, we used data on baseline thyrotropin stimulating hormone (TSH) levels and on dose of L-thyroxine used. Given that the majority of patients receiving relatively low doses of L-thyroxine have some endogenous thyroid function that could buffer any difference across preparations, we specifically investigated the comparative safety of brand vs. generic L-thyroxine among patients with thyroid cancer, patients receiving high doses of L-thyroxine, and patients with history of thyroid surgery.

## Methods

### Study design and data source

We conducted a retrospective analysis of de-identified administrative claims data linked with laboratory results from a large database, OptumLabs™ Data Warehouse (OLDW),

which includes commercially insured and Medicare Advantage enrollees throughout the United States [11]. The database contains longitudinal health information on enrollees and patients, representing a diverse mix of ages, ethnicities, and geographical regions across the United States [12]. The health plans provide comprehensive full insurance coverages for physician, hospital, and prescription drug services. Pharmacy claims include information on medications dispensed, size, and dates of prescriptions. Study data were accessed using techniques compliant with the Health Insurance Portability and Accountability Act of 1996. Because this study involved analysis of pre-existing, de-identified data, the Mayo Clinic Institutional Review Board declared it exempt from board approval.

### Study population

We included adults (≥ 18 years) who newly initiated either brand or generic L-thyroxine preparations between Jan 1, 2008 and Dec 31, 2018. We limited the study to patients who had continuous medical and pharmacy benefits for at least 12 months prior to the first L-thyroxine fill. Date of this prescription fill was considered the index date. At baseline, we excluded adults who were pregnant, had diagnosed hypopituitarism, hyperthyroidism, and those who had a medical condition or used medications that could affect thyroid stimulating hormone (TSH) levels (List of medications in Appendix Table #1) within 1 year before the index date [13]. Finally, we excluded adults who filled other forms of non-recommended thyroid replacement therapy before index date [13], including thyroid extracts, or T3 therapy, such as liothyronine, thyroid desiccated/extracts, as well as the following brands: Cytomel, Armour Thyroid, or Nature Thyroid.

### Exposure

We had a total of 33,038,162 fills for L-thyroxine during the study period, of which 2,178,862 were first fills (index date). For patients initiating L-thyroxine, we characterized whether the pharmacy fill was for a brand-name or generic L-thyroxine using First Databank. First Databank categorizes pharmacy products as generic if they are sold under a generic pharmacy label. Participants who switched across formulation (e.g., brand to generic) or switched to other thyroid hormones, disenrolled from the health plan, or ended treatment (defined as not refilling a prescription within 30 days of the end of the last filled prescription) during the study time were censored.

### Outcomes

Our primary outcomes were: atrial fibrillation (Afib), myocardial infarction (MI), congestive heart failure (CHF), pneumonia, falls, and spine and hip fractures. Afib and CHF outcomes were defined as primary diagnoses of Afib or CHF during hospitalization; MI and stroke outcomes were defined based on previous algorithm as primary or first secondary diagnosis of MI or stroke during hospitalization. Pneumonia outcome was defined as primary diagnosis of pneumonia during hospitalization. Falls, spine and hip fractures were defined as primary diagnosis in hospitalization or emergency department visits. We chose pneumonia hospitalization as the negative control because it has no known association with the exposure of interest [14]. If we were to find an association between rate of pneumonia and use of brand-name L-thyroxine, it would suggest that there are still important unadjusted confounding factors.

## Baseline characteristics

Baseline patient characteristics included age, sex, race/ethnicity, household income, census region, provider specialty (endocrinologist vs. other specialists), year of index prescription, L-thyroxine dose, Charlson comorbidity index [15], and conditions and medications that may be associated with the exposure and/or affect the risk of developing any of the outcomes of interest (Table 1). We defined comorbidities using International Classification Disease [ICD] billing codes (ICD-9-CM/ICD-10) from administrative claims.

## Statistical analyses

We used propensity score to match patients who initiated brand name L-thyroxine with those who initiated generic L-thyroxine [16]. Logistic regression model was used to estimate the propensity to receive brand name versus generic L-thyroxine. Based on clinical relevance and evidence from prior studies, the covariates related to the propensity score included demographics, baseline comorbidities, medications, and baseline TSH values shown in Table 1. We evaluated the balance among the treatment groups by comparing standardized mean differences of baseline covariates between the groups. A baseline characteristic was considered balanced if the maximum standardized mean difference was under 0.1. Patients on generic L-thyroxine were matched 1 to 1 to patients receiving brand L-thyroxine using nearest neighbor matching with a caliper of 0.2. Cox proportional hazards regression was used to compare the risk of experiencing the first incidence of the outcomes between brand versus generic L-thyroxine users in the matched cohort. The proportional hazard assumption was tested on the basis of Schoenfeld residuals and it was valid for all outcomes. Robust sandwich estimates were included to account for clustering within matched sets. The event rate, hazard ratio (HR), and 95% CI for each outcome of interest were calculated.

We then performed stratified analyses by age (under 65 and 65 years or older), history of thyroid cancer (present/absent), history of total thyroidectomy (present/absent), and dose of L-thyroxine used (under 100 and 100 mcg and over). Finally, we conducted a sensitivity analysis for primary outcomes after excluded patients taking L-thyroxine capsule, due to concerns that L-thyroxine capsules may have different absorption profile than L-thyroxine tablets and thus an impact on outcome of interest. Differences in hazard ratios (HR) by subgroups of interest were tested using interaction terms. Analyses were performed in SAS 9.4 (SAS Institute Inc., Cary, NC) and Stata version 14.1 (StatCorp, College Station, TX).  $P < 0.05$  was considered statistically significant.

## Results

The characteristics of the 195,046 participants who filled L-thyroxine between Jan 1, 2008 and Dec 31, 2018 are shown in Table 1. Most adults in the cohort were women (75%), 95% were under age 65, 64% were white, and 38% had baseline TSH levels between 0.3 and 4.4 mIU/L. Patients who received brand-name compared with generic L-thyroxine were more likely to be female (81% vs. 74%), to be white (68% vs. 63%), to have a household income  $> \$200,000$  (17% vs. 10%), to have a claim for thyroid cancer (4% vs. 1%), hypothyroidism (59% vs. 46%), and to have an endocrinologist as the prescriber specialty (29% vs. 10%).

After 1:1 propensity score matching, there were 35,667 patients in each treatment group. There were no significant differences between brand and generic L-thyroxine cohorts with respect to all of the examined factors (standardized mean differences <0.1). Matched patients' mean and median follow-up time were 0.88 (SD 1.05) years and 0.49 years, respectively. The crude event rates per 100 person-years and corresponding HRs for the two groups are shown in Table 2. Among all the matched patients, there were no significant differences between adults treated with brand versus generic L-thyroxine in atrial fibrillation (HR 0.96, 95% confidence interval [CI] 0.58–1.60), myocardial infarction (HR 0.66, 0.39–1.14), congestive heart failure (HR 1.30, 0.78–2.16), stroke (0.72, 0.49–1.06), spine (HR 0.87, 0.38–1.99) and hip fractures (HR 0.86, 0.26–2.82), or fall outcomes (HR 1.02, 0.14–7.32). Hospitalization rates for pneumonia—a negative control—also did not differ between groups (HR 0.85, 0.61–1.19). Kaplan-Meier curves for each outcome are depicted in Fig. 1. Furthermore, there were no significant interactions in any subgroups for the outcomes of interest (Table 3). Finally, sensitivity analysis after excluding L-thyroxine capture did not show different results in primary outcomes (Table 4).

## Discussion

In this observational study of 195,046 adults who initiated L-thyroxine replacement therapy, we found, among 35,667 propensity-matched patients, no statistically significant differences in cardiovascular outcomes or in rates of falls and fractures associated with the use of brand versus generic L-thyroxine formulations, including across subgroups of interest: younger vs. older adults, and those with potential low endogenous thyroid production. As patients tend to take the same formulation without switching, these results are important as they reassure clinicians and patients that taking generic L-thyroxine persistently does not offer any disadvantage compared to brand L-thyroxine in regards to the patient-important outcomes. In doing so, these results add important information about rates of falls and fractures to the study published by Smallridge et al. [10], in which they found no difference in the incident cardiovascular event rates between brand and generic L-thyroxine for patients with hypothyroidism.

In our study, the majority of patients received low doses of L-thyroxine (average initiating L-thyroxine dose of 70 mcg), likely due to having mild forms of thyroid dysfunction, suggesting that the majority had remaining intact endogenous thyroid function. Yet, we thought it was important to explore the effect of brand L-thyroxine formulation in patients with low or no endogenous thyroid production, as one study suggested that, for a group of 20 children with congenital hypothyroidism, with low or non-endogenous thyroid production, TSH values were different between brand and generic L-thyroxine [17]. In our study, we identified important subgroups of patients likely to have low or no thyroid endogenous production: patients with no thyroid, patients with thyroid cancer, and patients receiving full replacement L-thyroxine doses. Although we had limited statistical power due to the small size of the subgroups, we found no risk differences in the outcomes of interest, suggesting similar safety of brand and generic formulations in these patients. In contrast to the study conducted in children with congenital hypothyroidism, our study did not assess the short term biochemical impact of L-thyroxine formulation; rather, we used a pragmatic

approach to understand how persistent use of brand vs generic L-thyroxine formulation impacts patient-important outcomes.

### **Implication for clinical practice and research**

Our results support the concept that there is little difference in patient-important outcomes with use of either brand or generic L-thyroxine. Yet, there are still other aspects about L-thyroxine formulation that need to be investigated. For instance, we did not take into account the effect of switching formulations (between brand and generic) and switching manufacturers within each formulation. Experts recommend maintaining the same prescription L-thyroxine preparation throughout the duration of treatment (brand or generic) [13]. However, there are several manufacturers of generic L-thyroxine and it is known that switching among the products made by these manufacturers occurs at the pharmacy level, with such changes not necessarily noted by the patient. Given that clinicians are often unaware when patients are switched to L-thyroxine products made by different manufacturers, experts recommend using brand L-thyroxine. Future research is needed to explore the effect of the frequency and direction of switching between L-thyroxine formulations on thyroid hormone levels, and patient-important outcomes.

### **Limitations**

Our study has several limitations. This is an observational study, in which residual confounding factors may exist. To overcome this limitation, we used a negative control, pneumonia. The fact that the rate of pneumonia was not associated with the interventions (brand or generic L-thyroxine) argues against, but does not completely eliminate, the presence of residual confounding. Furthermore, we were not able to capture thyroid hormone doses during treatment, such that the similar safety between brand and generic L-thyroxine formulations could be explained by L-thyroxine dose adjustments that occurred during the follow-up of these patients. Whether these adjustments occurred similarly in both groups is unknown. The database used in this study does not capture all thyroid test results for the included patients during study follow-up, limiting our ability to understand the impact of thyroid function levels over time on the outcomes of interest. However, we did not note any difference in safety for patients with thyroid cancer (these patients' dose adjustments at follow-up may be similar in both groups as they both need to maintain a similar thyroid hormone target to suppress tumor growth). Moreover, there was a short follow-up period and the majority of patients dropped (censored at switch or discontinuation of L-thyroxine formulation or disenrollment from the health plan) from the analysis in the first 2 years. This limits our confidence about the impact of brand-name versus generic L-thyroxine use on the outcomes, as it is possible that for some of the included outcomes, several years of follow-up might be necessary to detect differences. Yet, if there was any effect of brand-name versus generic L-thyroxine on thyroid hormone values, this would likely affect cardiovascular outcomes over the short term. For instance, high levels of thyroid hormone may lead to heart arrhythmias, and subsequently other cardiovascular outcomes. Finally, our sample size includes mostly patients with commercial health insurance so the generalizability of our findings to under-insured populations is unclear.

## Conclusions

Our study of the comparative safety of persistent use of brand vs generic L-thyroxine for patients with mild forms of thyroid dysfunction showed that these formulations do not differ with respect to clinically important outcomes such as cardiovascular events, falls or fracture. Further research is needed to clarify if these findings are consistent among patients who switched between L-thyroxine formulations made by different manufacturers during the course of treatment.

## Funding

This project was supported by the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) as part of a financial assistance award [Center of Excellence in Regulatory Science and Innovation grant to Yale University and Mayo Clinic, U01FD005938]. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by FDA/HHS, or the U.S. Government.

## Data availability

All data generated or analyzed during this study are included in this published article.

## Appendix

### Appendix

**Appendix Table 1**

---

Clinically relevant medications that may affect thyroid function

---

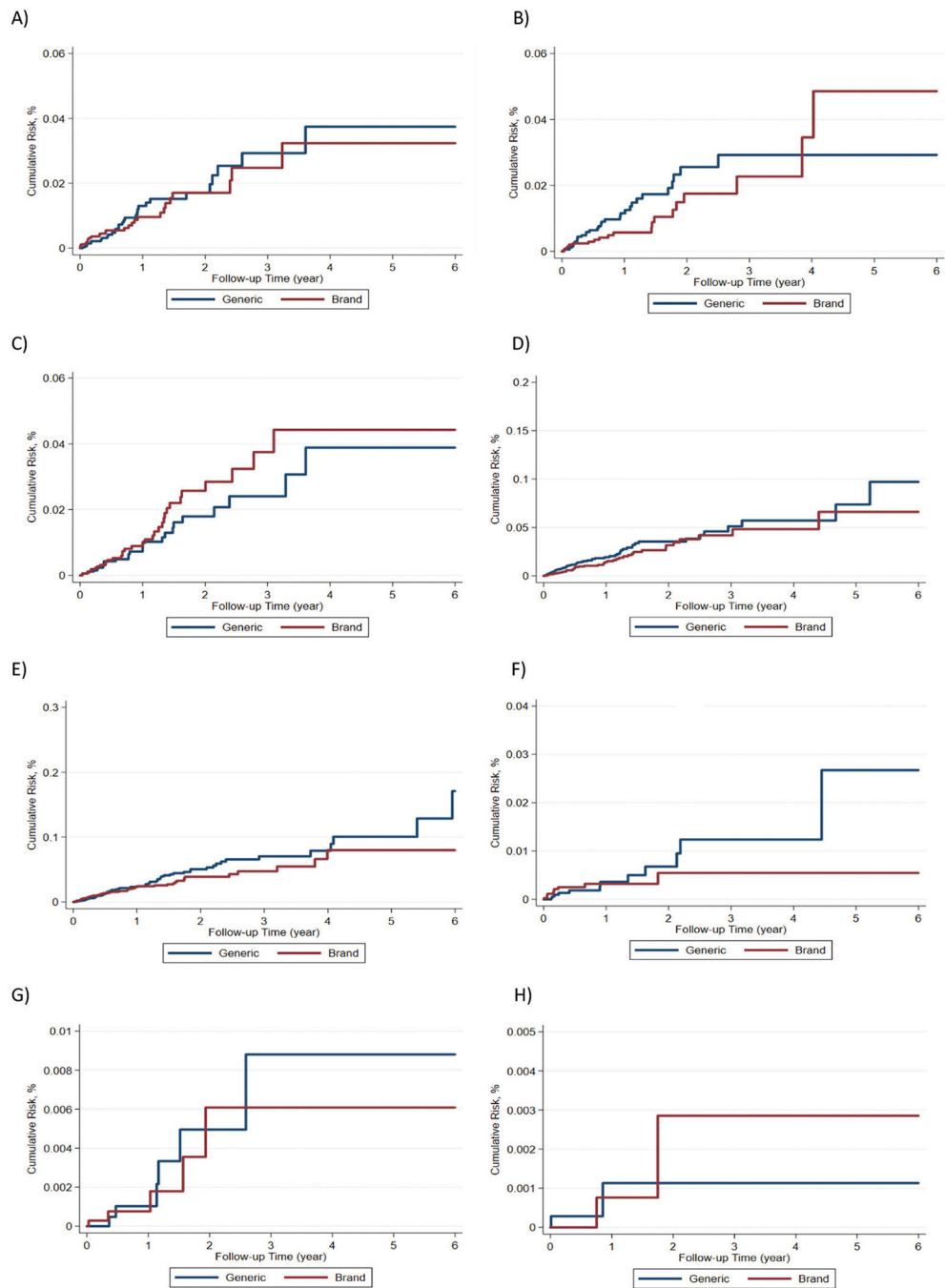
- Antithyroid Medications (PTU, Propylthiouracil, 6-N Propylthiouracil, Methimazole, Felimazole, Northyx, Tapazole, Thiamazole, Carbimazole, Benzylthiouracil, Methylthiouracil)
  - Lithium carbonate
  - Amiodarone hydrochloride
  - Phenytoin
  - Interferon alfa
  - Interleukin 2
  - Gefitinib
  - Erlotinib
  - Sorafenib
  - Sunitinib
  - Dasatinib
  - Lenvatinib
  - Imatinib
  - Cabozantinib
  - Vandetanib
- 

## References

1. Rodriguez-Gutierrez R, Maraka S, Ospina NS, Montori VM, Brito JP, Levothyroxine overuse: time for an about face? *Lancet Diab. Endocrinol* (2017). 10.1016/S2213-8587(16)30276-5



2. Kantor ED, Rehm CD, Haas JS, Chan AT, Giovannucci EL, Trends in prescription drug use among adults in the United States from 1999–2012. *Jama* (2015). 10.1001/jama.2015.13766
3. Viswanathan M, Golin CE, Jones CD et al. Interventions to improve adherence to self-administered medications for chronic diseases in the United States: a systematic review. *Ann. Intern. Med* (2012). 10.7326/0003-4819-157-11-201212040-00538
4. Ross JS, Rohde S, Sangaralingham L et al. Generic and brand-name thyroid hormone drug use among commercially insured and medicare beneficiaries, 2007 Through 2016. *J. Clin. Endocrinol. Metab* (2019). 10.1210/jc.2018-02197
5. Huo N, Chen L, Mishuk AU et al. Generic levothyroxine initiation and substitution among Medicare and Medicaid populations: a new user cohort study. *Endocrine* (2020). 10.1007/s12020-020-02211-w
6. American Thyroid Association, The Endocrine Society, and American Association of Clinical Endocrinologists. Joint Statement on the U.S. Food and Drug Administration’s Decision Regarding Bioequivalence of Levothyroxine Sodium (2021). <https://www.endocrine.org/advocacy/position-statements/bioequivalence-of-sodium-levothyroxine>
7. Endocrine Society. Bioequivalence of Sodium Levothyroxine (2008). [https://www.optum.com/content/dam/optum/resources/productSheets/5302\\_Data\\_Assets\\_Chart\\_Sheet\\_ISPOR.pdf](https://www.optum.com/content/dam/optum/resources/productSheets/5302_Data_Assets_Chart_Sheet_ISPOR.pdf)
8. Grossman A, Feldhamer I, Meyerovitch J, Treatment with levothyroxin in subclinical hypothyroidism is associated with increased mortality in the elderly. *Eur. J. Intern. Med* (2018). 10.1016/j.ejim.2017.11.010
9. Taylor PN, Iqbal A, Minassian C et al. Falling threshold for treatment of borderline elevated thyrotropin levels—balancing benefits and risks: evidence from a large community-based study. *JAMA Intern. Med* (2014). 10.1001/jamainternmed.2013.11312
10. Smallridge RC, Sangaralingham LR, Mwangi R, Kusumoto F, Van Houten H, Bernet V, Comparison of Incident Cardiovascular Event Rates Between Generic and Brand l-Thyroxine for the Treatment of Hypothyroidism. *Mayo Clinic Proceedings* (2019)
11. Wallace PJ, Shah ND, Dennen T, Bleicher PA, Crown WH, Optum Labs: building a novel node in the learning health care system. *Health Affairs* (2014). 10.1377/hlthaff.2014.0038
12. OptumLabs. Real world health care experiences from over 150 million unique individuals since 1993 (2015). [https://www.optum.com/content/dam/optum/resources/productSheets/5302\\_Data\\_Assets\\_Chart\\_Sheet\\_ISPOR.pdf](https://www.optum.com/content/dam/optum/resources/productSheets/5302_Data_Assets_Chart_Sheet_ISPOR.pdf)
13. Jonklaas J, Bianco A, Bauer A et al. , American Thyroid Association Task Force on Thyroid Hormone Replacement. Guidelines for the treatment of hypothyroidism: prepared by the american thyroid association task force on thyroid hormone replacement. *Thyroid* (2014). 10.1089/thy.2014.0028
14. Desai JR, Hyde CL, Kabadi S et al. Utilization of positive and negative controls to examine comorbid associations in observational database studies. *Med. Care* (2017). 10.1097/MLR.0000000000000640
15. Deyo RA, Cherkin DC, Ciol MA, Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J. Clin. Epidemiol* (1992). 10.1016/0895-4356(92)90133-8
16. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF, A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Stat. Med* (2013). 10.1002/sim.5753
17. Carswell JM, Gordon JH, Popovsky E, Hale A, Brown RS, Generic and brand-name L-thyroxine are not bioequivalent for children with severe congenital hypothyroidism. *J. Clin. Endocrinol. Metab* (2013). 10.1210/jc.2012-3125



**Fig. 1.** Kaplan–Meier curves for outcomes of interest; **a** atrial fibrillation, **b** myocardial infarction, **c** congestive heart failure, **d** stroke, **e** pneumonia, **f** spine fractures, **g** hip fractures, **h** falls

**Table 1**

Baseline characteristics of unmatched and matched new-user cohorts of levothyroxine

Characteristics	Pre-match		Post-match		Standardized mean difference	Standardized mean difference
	Generic T4 (N = 157,830)	Brand T4 (N = 37,216)	Generic T4 (N = 35,667)	Brand T4 (N = 35,667)		
<i>Initiating dose, µg/d</i>						
Mean (SD)	61.3 (38.8)	72.2 (42.8)	70.4 (43.4)	70.6 (41.8)	0.00	0.00
Median	50.0	50.0	50.0	50.0		
<i>TSH level, mIU/L</i>						
<0.3	5531 (3.5%)	2326 (6.3%)	1972 (5.5%)	2005 (5.6%)	0.00	0.00
0.3–4.4	55,858 (35.4%)	17,778 (47.8%)	16,864 (47.3%)	16,757 (47.0%)	0.01	0.01
4.5–9.9	77,699 (49.2%)	13,549 (36.4%)	13,420 (37.6%)	13,410 (37.6%)	0.00	0.00
10–19.9	10,811 (6.8%)	1923 (5.2%)	1843 (5.2%)	1902 (5.3%)	0.01	0.01
>19.9	7931 (5.0%)	1640 (4.4%)	1568 (4.4%)	1593 (4.5%)	0.00	0.00
<i>Age group, y</i>						
<35	30,419 (19.3%)	7477 (20.1%)	7335 (20.6%)	7169 (20.1%)	0.01	0.01
35–44	33,649 (21.3%)	8621 (23.2%)	8369 (23.5%)	8199 (23.0%)	0.01	0.01
45–54	44,419 (28.1%)	10,925 (29.4%)	10,267 (28.8%)	10,425 (29.2%)	0.01	0.01
55–64	40,747 (25.8%)	8857 (23.8%)	8428 (23.6%)	8559 (24.0%)	0.01	0.01
65+	8596 (5.4%)	1336 (3.6%)	1268 (3.6%)	1315 (3.7%)	0.01	0.01
<i>Gender</i>						
Female	116,083 (73.5%)	30,157 (81.0%)	28,889 (81.0%)	28,762 (80.6%)	0.01	0.01
Male	41,747 (26.5%)	7059 (19.0%)	6778 (19.0%)	6905 (19.4%)	0.01	0.01
<i>Race/ethnicity</i>						
White	98,726 (62.6%)	25,232 (67.8%)	24,126 (67.6%)	24,007 (67.3%)	0.01	0.01
Black	17,710 (11.2%)	3733 (10.0%)	3584 (10.0%)	3633 (10.2%)	0.01	0.01
Hispanic	23,085 (14.6%)	4347 (11.7%)	4202 (11.8%)	4239 (11.9%)	0.00	0.00
Asian	12,491 (7.9%)	2449 (6.6%)	2447 (6.9%)	2401 (6.7%)	0.01	0.01
Unknown	5818 (3.7%)	1455 (3.9%)	1308 (3.7%)	1387 (3.9%)	0.01	0.01
<i>Patients census region</i>						
Midwest	17,463 (11.1%)	3077 (8.3%)	2902 (8.1%)	3025 (8.5%)	0.01	0.01

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Characteristics	Pre-match			Post-match		
	Generic T4 (N = 157,830)	Brand T4 (N = 37,216)	Standardized mean difference	Generic T4 (N = 35,667)	Brand T4 (N = 35,667)	Standardized mean difference
Northeast	12,895 (8.2%)	3325 (8.9%)	0.03	3174 (8.9%)	3185 (8.9%)	0.00
South	93,116 (59.0%)	24,542 (65.9%)	0.14	23,529 (66.0%)	23,309 (65.4%)	0.01
West	34,162 (21.6%)	6242 (16.8%)	0.12	6027 (16.9%)	6118 (17.2%)	0.01
Unknown	194 (0.1%)	30 (0.1%)	0.01	35 (0.1%)	30 (0.1%)	0.01
<i>Charlson group</i>						
0	127,786 (81.0%)	30,529 (82.0%)	0.03	29,594 (83.0%)	29,398 (82.4%)	0.02
1	18,339 (11.6%)	3843 (10.3%)	0.04	3591 (10.1%)	3733 (10.50%)	0.01
2	6699 (4.2%)	1910 (5.1%)	0.04	1648 (4.6%)	1659 (4.7%)	0.00
3+	5006 (3.2%)	934 (2.5%)	0.04	834 (2.30%)	877 (2.50%)	0.00
<i>Household income</i>						
<\$40,000	22,700 (14.4%)	3434 (9.2%)	0.16	3339 (9.4%)	3411 (9.6%)	0.01
\$40,000–\$74,999	37,782 (23.9%)	7170 (19.3%)	0.11	7014 (19.7%)	7043 (19.7%)	0.00
\$75,000–\$124,999	44,218 (28.0%)	10,273 (27.6%)	0.01	9999 (28.0%)	9967 (27.9%)	0.00
\$125,000–\$199,999	25,064 (15.9%)	7623 (20.5%)	0.12	7199 (20.2%)	7200 (20.2%)	0.00
\$200,000	16,263 (10.3%)	6455 (17.3%)	0.21	5934 (16.6%)	5851 (16.4%)	0.01
Unknown	11,803 (7.5%)	2261 (6.1%)	0.06	2182 (6.1%)	2195 (6.2%)	0.00
<i>Year of index prescription</i>						
2008	7999 (5.1%)	3938 (10.6%)	0.21	3519 (9.9%)	3480 (9.8%)	0.00
2009	11,520 (7.3%)	4591 (12.3%)	0.17	4393 (12.3%)	4225 (11.8%)	0.01
2010	13,139 (8.3%)	4777 (12.8%)	0.15	4435 (12.4%)	4478 (12.6%)	0.00
2011	12,090 (7.7%)	3751 (10.1%)	0.09	3501 (9.8%)	3565 (10.0%)	0.01
2012	12,529 (7.9%)	3189 (8.6%)	0.02	3215 (9.0%)	3116 (8.7%)	0.01
2013	15,707 (10.0%)	3661 (9.8%)	0.00	3595 (10.1%)	3581 (10.0%)	0.00
2014	14,988 (9.5%)	2897 (7.8%)	0.06	2960 (8.3%)	2856 (8.0%)	0.01
2015	16,706 (10.6%)	2859 (7.7%)	0.10	2928 (8.2%)	2835 (7.9%)	0.01
2016	21,397 (13.6%)	3224 (8.7%)	0.16	3313 (9.3%)	3208 (9.0%)	0.01
2017	18,707 (11.9%)	2940 (7.9%)	0.13	2552 (7.2%)	2934 (8.2%)	0.04
2018	13,048 (8.3%)	1389 (3.7%)	0.19	1256 (3.5%)	1389 (3.9%)	0.02
<i>Prescriber specialty</i>						

Characteristics	Pre-match			Post-match		
	Generic T4 (N = 157,830)	Brand T4 (N = 37,216)	Standardized mean difference	Generic T4 (N = 35,667)	Brand T4 (N = 35,667)	Standardized mean difference
	Endocrinology	15,193 (9.6%)	10,770 (28.9%)	0.51	8983 (25.2%)	9232 (25.9%)
Primary care	96,776 (61.3%)	18,179 (48.8%)	0.25	18,315 (51.3%)	18,173 (51.0%)	0.01
Other	33,765 (21.4%)	5902 (15.9%)	0.14	6028 (16.9%)	5901 (16.5%)	0.01
Missing/Unknown	12,096 (7.7%)	2365 (6.4%)	0.05	2341 (6.6%)	2361 (6.6%)	0.00
<i>Baseline conditions</i>						
Myocardial infarction	482 (0.3%)	67 (0.2%)	0.03	61 (0.2%)	66 (0.2%)	0.00
Congestive heart failure	2407 (1.5%)	352 (0.9%)	0.05	330 (0.9%)	345 (1.0%)	0.00
Peripheral vascular disease	1962 (1.2%)	349 (0.9%)	0.03	321 (0.9%)	337 (0.9%)	0.01
Cerebrovascular disease	2093 (1.3%)	399 (1.1%)	0.02	364 (1.0%)	390 (1.1%)	0.01
Cardiac arrhythmia	7333 (4.6%)	1527 (4.1%)	0.03	1448 (4.1%)	1467 (4.1%)	0.00
Hypothyroidism	72,057 (45.7%)	21,931 (58.9%)	0.27	20,592 (57.7%)	20,606 (57.8%)	0.00
Hypertension	45,598 (28.9%)	8251 (22.2%)	0.16	7861 (22.0%)	8072 (22.6%)	0.01
Anemia	6716 (4.3%)	1596 (4.3%)	0.00	1522 (4.3%)	1554 (4.4%)	0.00
Diabetes	11,740 (7.4%)	2286 (6.1%)	0.05	2067 (5.8%)	2221 (6.2%)	0.02
Obesity	17,365 (11.0%)	2894 (7.8%)	0.11	2668 (7.5%)	2861 (8.0%)	0.02
Weight loss	2053 (1.3%)	419 (1.1%)	0.02	414 (1.2%)	411 (1.2%)	0.00
Depression	18510 (11.7%)	4093 (11.0%)	0.02	3861 (10.8%)	3979 (11.2%)	0.01
Psychoses	728 (0.5%)	131 (0.4%)	0.02	120 (0.3%)	129 (0.4%)	0.00
Thyroid cancer	2125 (1.3%)	1406 (3.8%)	0.15	1075 (3.0%)	1033 (2.9%)	0.01
Paget's	*	*	0.00	*	*	0.00
Vitamin D deficiency	19,785 (12.5%)	5606 (15.1%)	0.07	5190 (14.6%)	5267 (14.8%)	0.01
Osteoporosis	1583 (1.0%)	450 (1.2%)	0.02	417 (1.2%)	424 (1.2%)	0.00
Osteopenia	3763 (2.4%)	1387 (3.7%)	0.08	1249 (3.5%)	1274 (3.6%)	0.00
Hyperparathyroidism	489 (0.3%)	235 (0.6%)	0.05	210 (0.6%)	193 (0.5%)	0.01
Cushing's syndrome	39 (0.0%)	13 (0.0%)	0.01	12 (0.0%)	13 (0.0%)	0.00
Cirrhosis	2733 (1.7%)	496 (1.3%)	0.03	448 (1.3%)	484 (1.4%)	0.01
Hypogonadism	1321 (0.8%)	188 (0.5%)	0.04	183 (0.5%)	187 (0.5%)	0.00
Rheumatoid arthritis	4686 (3.0%)	1103 (3.0%)	0.00	1029 (2.9%)	1063 (3.0%)	0.01
Menopause	8510 (5.4%)	2973 (8.0%)	0.10	2765 (7.8%)	2774 (7.8%)	0.00

Characteristics	Pre-match			Post-match		
	Generic T4 (N = 157,830)	Brand T4 (N = 37,216)	Standardized mean difference	Generic T4 (N = 35,667)	Brand T4 (N = 35,667)	Standardized mean difference
Alzheimer's disease	99 (0.1%)	*	0.02	*	*	0.00
Dementia	260 (0.2%)	48 (0.1%)	0.01	38 (0.1%)	48 (0.1%)	0.01
Parkinson's disease	63 (0.0%)	*	0.01	*	*	0.00
Bulimia or anorexia	293 (0.2%)	86 (0.2%)	0.01	76 (0.2%)	81 (0.2%)	0.00
Vertebral	214 (0.1%)	47 (0.1%)	0.00	40 (0.1%)	45 (0.1%)	0.00
Hip	65 (0.0%)	16 (0.0%)	0.00	15 (0.0%)	14 (0.0%)	0.00
Fracture	338 (0.2%)	63 (0.2%)	0.01	60 (0.2%)	60 (0.2%)	0.00
Frailty	2190 (1.4%)	381 (1.0%)	0.03	336 (0.9%)	367 (1.0%)	0.01
Pacemaker device	657 (0.4%)	109 (0.3%)	0.02	100 (0.3%)	106 (0.3%)	0.00
COPD	8048 (5.1%)	1680 (4.5%)	0.03	1586 (4.4%)	1613 (4.5%)	0.00
Pneumonia	2208 (1.4%)	377 (1.0%)	0.04	346 (1.0%)	363 (1.0%)	0.01
Stroke	1009 (0.6%)	147 (0.4%)	0.03	123 (0.3%)	145 (0.4%)	0.01
Thyroid surgery	2783 (1.8%)	1079 (2.9%)	0.08	997 (2.8%)	971 (2.7%)	0.00
<i>Medications (180 days prior)</i>						
ACEi	14,066 (8.9%)	2186 (5.9%)	0.12	2067 (5.8%)	2154 (6.0%)	0.01
ARB	10,110 (6.4%)	1943 (5.2%)	0.05	1808 (5.1%)	1882 (5.3%)	0.01
β-blockers	12,652 (8.0%)	2304 (6.2%)	0.07	2210 (6.2%)	2237 (6.3%)	0.00
Loop diuretics	3300 (2.1%)	486 (1.3%)	0.06	456 (1.3%)	481 (1.3%)	0.01
Aldosterone antagonist	1776 (1.1%)	404 (1.1%)	0.00	380 (1.1%)	388 (1.1%)	0.00
Digoxin	328 (0.2%)	61 (0.2%)	0.01	58 (0.2%)	59 (0.2%)	0.00
Statin	21,987 (13.9%)	4188 (11.3%)	0.08	4019 (11.3%)	4073 (11.4%)	0.01
ACEi/ARB	23,589 (14.9%)	4024 (10.8%)	0.12	3769 (10.6%)	3932 (11.0%)	0.02
Osteoporosis treatment	14,865 (9.4%)	3193 (8.6%)	0.03	2992 (8.4%)	3073 (8.6%)	0.01
Estrogen	7124 (4.5%)	2074 (5.6%)	0.05	1985 (5.6%)	1980 (5.6%)	0.00

\* Cannot include N when cell size is <11 to protect patient confidentiality

**Table 2**

Event rates and hazards ratios (95% confidence interval) of study outcomes in propensity score-matched cohorts of new users of brand or generic levothyroxine

Outcomes	Events per 1000 patient years*		Hazard ratio (95% CI) brand vs. generic	P-value
	Brand T4	Generic T4		
Atrial fibrillation	0.95	0.98	0.96 (0.58, 1.60)	0.88
Myocardial infarction	0.72	1.08	0.66 (0.39, 1.14)	0.14
Congestive heart failure	1.11	0.85	1.30 (0.78, 2.16)	0.31
Stroke	1.44	1.98	0.72 (0.49, 1.06)	0.10
Pneumonia	2.06	2.41	0.85 (0.61, 1.19)	0.34
Spine fractures	0.33	0.37	0.87 (0.38, 1.99)	0.73
Hip fractures	0.16	0.19	0.86 (0.26, 2.82)	0.81
Falls	0.07	0.06	1.02 (0.14, 7.32)	0.99

\* Censored on end of coverage, 30 days after end of treatment, event or brand to generic, generic to brand switch or switch to other thyroid medications

**Table 3**

Subgroup sensitivity analyses

Outcomes	Subgroup	Events per 1000 patient years*		Hazard ratio (95% CI) brand vs. generic	P-value for interaction
		Brand T4	Generic T4		
Atrial fibrillation		0.95	0.98	0.96 (0.58, 1.60)	
	Age < 65 (N = 68,751)	0.68	0.81	0.82 (0.46, 1.48)	0.33
	Age 65 (N = 2583)	8.80	5.88	1.47 (0.52, 4.13)	
	With thyroid cancer (N = 2108)	1.03	1.06	1.02 (0.06, 16.01)	1.00
	Without thyroid cancer (N = 69,226)	0.95	0.97	0.96 (0.57, 1.61)	
	Dose < 100 mcg (N = 51,611)	0.61	0.62	0.98 (0.46, 2.08)	0.92
Myocardial infarction	Dose 100 mcg (N = 19,723)	1.74	1.87	0.92 (0.46, 1.82)	
	With thyroid surgery (N = 1968)	0	1.54	NA	
	Without thyroid surgery (N = 69,366)	0.97	0.96	0.99 (0.60, 1.66)	
	Age < 65	0.72	1.08	0.66 (0.39, 1.14)	
	Age 65	0.68	0.88	0.76 (0.43, 1.37)	0.23
	With thyroid cancer	1.93	7.09	0.27 (0.06, 1.29)	0.81
Congestive heart failure	Without thyroid cancer	1.04	1.13	0.85 (0.05, 14.20)	
	Dose < 100 mcg	0.71	1.07	0.65 (0.38, 1.14)	
	Dose 100 mcg	0.51	0.84	0.61 (0.29, 1.28)	0.75
	With thyroid surgery	1.20	1.66	0.71 (0.33, 1.56)	
	Without thyroid surgery	1.59	1.66	0.94 (0.05, 16.08)	0.79
	Age < 65	0.70	1.06	0.65 (0.38, 1.13)	
	Age 65	1.11	0.85	1.30 (0.78, 2.16)	
	With thyroid cancer	0.91	0.75	1.22 (0.70, 2.12)	0.63
	Without thyroid cancer	6.77	3.97	1.71 (0.51, 5.77)	
	Dose < 100 mcg	0	0	NA	0.32
	Dose 100 mcg	1.15	0.87	1.31 (0.79, 2.17)	
	With thyroid surgery	0.98	0.83	1.17 (0.63, 2.18)	0.58
Without thyroid surgery	1.41	0.88	1.59 (0.66, 3.81)		
	0	0	NA		
	1.14	0.86	1.30 (0.79, 2.16)		



Outcomes	Subgroup	Events per 1000 patient years*		Hazard ratio (95% CI) brand vs. generic	P-value for interaction
		Brand T4	Generic T4		
Stroke		1.44	1.98	0.72 (0.49, 1.06)	
	Age < 65	1.29	1.66	0.77 (0.51, 1.17)	0.39
	Age 65	5.79	11.95	0.48 (0.18, 1.28)	
	With thyroid cancer	2.07	2.33	0.90 (0.13, 6.42)	0.83
	Without thyroid cancer	1.42	1.97	0.71 (0.48, 1.06)	
	Dose < 100 mcg	1.36	1.76	0.77 (0.47, 1.24)	0.67
	Dose 100 mcg	1.63	2.54	0.64 (0.33, 1.22)	
	With thyroid surgery	4.73	1.63	2.92 (0.30, 28.73)	0.21
	Without thyroid surgery	1.37	1.99	0.68 (0.46, 1.01)	
			2.06	2.41	0.85 (0.61, 1.19)
Pneumonia	Age < 65	1.69	2.23	0.75 (0.52, 1.08)	0.11
	Age 65	12.63	7.71	1.66 (0.69, 3.99)	
	With thyroid cancer	3.11	2.17	1.44 (0.23, 9.02)	0.56
	Without thyroid cancer	2.03	2.41	0.83 (0.59, 1.17)	
	Dose < 100 mcg	1.87	2.20	0.85 (0.56, 1.28)	0.98
	Dose 100 mcg	2.50	2.92	0.86 (0.49, 1.50)	
	With thyroid surgery	4.72	3.32	1.54 (0.24, 9.79)	0.56
	Without thyroid surgery	2.01	2.39	0.83 (0.59, 1.17)	
		0.33	0.37	0.87 (0.38, 1.99)	
		0.30	0.35	0.85 (0.35, 2.03)	0.92
Spine fractures	Age < 65	0.97	0.97	1.00 (0.06, 15.53)	
	Age 65	0	0	NA	
	With thyroid cancer	0.34	0.38	0.87 (0.38, 2.00)	
	Without thyroid cancer	0.47	0.39	1.18 (0.48, 2.85)	
	Dose < 100 mcg	0	0.32	1.86E-16 (6.00E-17, 5.78E-16)	
	Dose 100 mcg	0	0	NA	
	With thyroid surgery	0.33	0.38	0.86 (0.38, 1.99)	
	Without thyroid surgery	0.16	0.19	0.86 (0.26, 2.82)	
	Age < 65	0.14	0.16	0.83 (0.22, 3.09)	0.92
	Age 65	0.97	1.00	0.94 (0.06, 14.50)	
Hip fractures					

Outcomes	Subgroup	Events per 1000 patient years*		Hazard ratio (95% CI) brand vs. generic	P-value for interaction	
		Brand T4	Generic T4			
Falls	With thyroid cancer	0	0	NA		
	Without thyroid cancer	0.17	0.19	0.87 (0.27, 2.83)		
	Dose < 100 mcg	0.19	0.26	0.71 (0.20, 2.47)		
	Dose 100 mcg	0.11	0	7.02E + 14 (9.88E + 13, 4.98E + 15)		
	With thyroid surgery	0	0	NA		
	Without thyroid surgery	0.17	0.19	0.87 (0.26, 2.82)		
			0.07	0.06	1.02 (0.14, 7.32)	
	Age < 65	0.07	0.06	1.02 (0.14, 7.35)	0.98	
	Age 65	0	0	NA		
	With thyroid cancer	0	0	NA		
	Without thyroid cancer	0.07	0.06	1.02 (0.14, 7.37)		
	Dose < 100 mcg	0.09	0.09	1.04 (0.14, 7.57)	0.97	
	Dose 100 mcg	0	0	NA		
	With thyroid surgery	0	0	NA		
Without thyroid surgery	0.07	0.06	1.02 (0.14, 7.34)			

\* Censored on end of coverage, 30 days after end of treatment, event or brand to generic, generic to brand switch or switch to other thyroid medications

**Table 4**

Sensitivity analysis for primary outcomes after dropping L-thyroxine capsule

Outcomes	Events per 1000 patient years*		Hazard ratio (95% CI) brand vs. generic	P-value
	Brand T4	Generic T4		
Atrial fibrillation	0.93	0.98	0.94 (0.56, 1.57)	0.82
Myocardial infarction	0.70	1.08	0.64 (0.37, 1.10)	0.11
Congestive heart failure	1.13	0.85	1.32 (0.80, 2.18)	0.28
Stroke	1.46	1.98	0.73 (0.50, 1.08)	0.11
Pneumonia	2.06	2.41	0.85 (0.61, 1.19)	0.34
Spine fractures	0.33	0.37	0.88 (0.38, 2.02)	0.76
Hip fractures	0.17	0.19	0.87 (0.27, 2.85)	0.82
Falls	0.07	0.06	1.03 (0.14, 7.40)	0.98

\* Event rates and hazards ratios (95% confidence interval) of study outcomes in propensity score-matched cohorts of new users of brand or generic levothyroxine

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript