The Tube Versus Trabeculectomy IRIS® Registry Study: Cohort Selection and Follow-up and Comparisons to the Randomized Controlled Trial

ELIZABETH A. VANNER, CATHERINE Q. SUN, MATTHEW J. MCSOLEY, PATRICE J. PERSAD, WILLIAM J. FEUER, FLORA LUM, SCOTT P. KELLY, RICHARD K. PARRISH, TA C. CHANG, AND STEVEN J. GEDDE

• PURPOSE: To assess the feasibility of replicating a randomized controlled trial (RCT) with a cohort of eyes, from IRIS® Registry data, analogous to the Tube Versus Trabeculectomy (TVT) RCT cohort and compare characteristics and follow-up.

• DESIGN: Comparison of RCT and IRIS Registry cohorts and follow-up.

• METHODS: We identified a cohort of IRIS Registry eyes (2013-2017) that received either a glaucoma drainage implant (tube) or trabeculectomy after a previous trabeculectomy and/or cataract extraction; extracted clinical and demographic characteristics for baseline surgery and follow-up visits through 1 year; and compared treatment groups in the IRIS Registry cohort and this cohort to the TVT RCT cohort.

• RESULTS: The IRIS Registry cohort included 419 eyes: 183 (43.7%) trabeculectomy; 236 (56.3%) tube. There were significant differences between treatment groups, including race (White: trabeculectomy 61.8%, tube 44.9%; Black: trabeculectomy 20.8%, tube 35.6%; $P = .003$) and the percentage of follow-up visits completed (trabeculectomy 88.4%, tube 83.8%, $P = .004$). There were also significant differences between the TVT IRIS Registry cohort and the TVT RCT cohort in the percentage of follow-up visits completed (IRIS Registry 85.6%, RCT 96.1%, $P < .001$) and in the probability of having a 1-year follow-up visit (IRIS Registry 81.4%, RCT 89.2%, $P = .011$).

• CONCLUSION: The TVT IRIS Registry cohort had several significant treatment group differences at baseline, whereas there had been none in the TVT RCT cohort. Follow-up in the TVT IRIS Registry cohort was inferior to that of the TVT RCT. Some data needed to refine the selection of eyes for the cohort were not available in the IRIS Registry. (Am J Ophthalmol 2021;224:43–52. © 2020 Elsevier Inc. All rights reserved.)

TWO ISSUES CURRENTLY UNDER DEBATE ARE TO what extent electronic health record (EHR) data can be used to compare health care outcomes among treatment alternatives and whether this can replace the need for randomized controlled trials (RCT). Before replacing RCTs is even considered, the ability of EHR data to confirm the results of existing RCTs must be thoroughly explored. The difficulty of even replicating the patient cohorts of RCTs with EHR data was explored recently, with the conclusion that only 15.0% of the cohorts from 220 US-based trials "could be replicated using observational data because their intervention, indication, inclusion and exclusion criteria, and primary end points could be routinely ascertained from insurance claims and/or EHR data."[1] In addition, Averitt and associates assessed whether applying the inclusion and exclusion criteria from an RCT to EHR data would result in selecting a real-world data (RWD) cohort that was similar to the RCT cohort, but found that this was not feasible.[2]

On June 23, 2020, a PUBMED search for "IRIS Registry" returned 17 results, but no results for "IRIS Registry" and Randomized Control (or Controlled) Trial. On that date, a search for "big data" and Randomized Control (or Controlled) Trial returned 133 results. Although some of these 133 results discussed clinical decision-making based on RCT data vs "big data," no direct comparisons of a specific RCT to an analogous cohort from EHR data were found.

The overall purpose for this research was to assess the feasibility of using retrospective, EHR-based data to replicate an RCT. Specifically, this research was supported by a 2018 Research to Prevent Blindness/American Academy of Ophthalmology Award for IRIS® Registry (Intelligent Research in Sight) Research to compare 1-year postsurgical results from the Tube Versus Trabeculectomy (TVT) RCT with an IRIS Registry cohort. The TVT RCT compared the safety and efficacy of a Baerveldt 350 glaucoma implant

AJO.com Supplemental Material available at AJO.com.
Accepted for publication Nov 27, 2020.

From the Bascom Palmer Eye Institute, Miami, Florida, USA (E.A.V., C.Q.S., M.J.M., P.J.P., W.J.F., R.K.P., T.C.C., S.J.G.); University of California, San Francisco, California, USA (C.Q.S.); University of Miami Miller School of Medicine, Miami, Florida, USA (M.J.M.); and American Academy of Ophthalmology, San Francisco, California, USA (F.L., S.K.).

Inquiries to Elizabeth A. Vanner, Director of Biostatistics, Bascom Palmer Eye Institute; 1501 NW 9th Avenue, Room 335, Miami, FL 33136, USA; e-mail: EAV45@med.miami.edu

0002-9394/$36.00
https://doi.org/10.1016/j.ajo.2020.11.014

(now Johnson & Johnson Vision, Milpitas, California, USA) to trabeculectomy with mitomycin C "in eyes that had undergone previous filtering surgery, cataract surgery with intraocular lens implantation, or both."[3] Here we describe the TVT IRIS Registry cohort creation, variables for which the TVT IRIS Registry treatment groups differed at baseline, and how this cohort differed from the TVT RCT cohort at baseline and during follow-up.

METHODS

• **DESIGN:** This study was designed to assess the feasibility of replicating an RCT with EHR data by comparing the patient cohort and results from the TVT RCT to those from a cohort of analogous eyes from IRIS Registry data. Because IRIS Registry data are deidentified, no informed consent was required, and the University of Miami Institutional Review Board approved this study as exempt. The study and data accumulation were in conformity with all country, federal, or state laws and adhered to the tenets of the Declaration of Helsinki and the Health Insurance Portability and Accountability Act. The TVT RCT is registered at http://www.clinicaltrials.gov (NCT00306852).

• **SETTING, PATIENTS, AND STUDY POPULATION:** This study, conducted at the Bascom Palmer Eye Institute, Miami, Florida, USA, used data from the American Academy of Ophthalmology (Academy) IRIS Registry, which is the first comprehensive ophthalmology database in the United States. These data are collected from the EHR systems of participating providers, and the IRIS Registry amalgamates data from all providers for each patient.[4] As of January 1, 2018, 14,245 physicians (ophthalmologists plus eligible clinicians) from 2,903 practices had signed up for EHR integration, and the IRIS Registry database contained 182.68 million visits for 44.23 million unique patients. Our initial data from the IRIS Registry included all eyes that had a glaucoma drainage implant/aqueous shunt ("tube") and/or a trabeculectomy from 2013 through 2017.

• **DATA:** The following data types were used in this project: demographic, visual acuity (VA), intraocular pressure (IOP), diagnosis, procedure, and medication. A list of all variables is found in Supplemental Table 1 (Supplemental Material available at AJO.com). All service date data were limited to year and week. We classified medications as being for glaucoma, ophthalmic use, diabetes, or hypertension. Medication data were not eye specific, so all medications for a patient were assumed to apply to the study eye. A list of the International Classification of Diseases (ICD) diagnosis and Current Procedural Terminology (CPT) procedure codes that were used to select eyes and to determine risk factors and outcomes may be found in Supplemental Table 2 (Supplemental Material available at AJO.com). Glaucoma types were classified using ICD diagnosis codes. Patients were classified as having diabetes and/or hypertension using diagnoses and medications. Procedures were classified using CPT codes such as tube (66179 or 66180), trabeculectomy (66172 and 66170), cataract extraction (66982 or 66984), and other relevant procedures used in the TVT RCT.

• **INCLUSION/EXCLUSION CRITERIA:** Based on the inclusion criteria of the TVT RCT,[3] we identified an initial cohort of eyes that had 2 surgeries: a "baseline" surgery (tube or trabeculectomy) after having a "qualifying" surgery (previous trabeculectomy and/or cataract surgery), both of which must have been specified for the same eye. From this initial cohort, we eliminated eyes (1) with less than 1 year of data in the IRIS Registry before the "qualifying" surgery, (2) with fewer than 90 days between the "qualifying" and "baseline" surgeries, (3) without 1 or more data types (Supplemental Table 1), (4) that did not pass the TVT RCT inclusion/exclusion screen,[3] and (5) that had their baseline surgery in 2017, because most of these lacked 1-year follow-up data. Only the first qualifying eye of a patient was included. The TVT RCT inclusion and exclusion criteria,[3,5,6] listed in Supplemental Table 3 (Supplemental Material available at AJO.com), were assessed using both diagnosis and procedure codes, although, as indicated, it was not always possible to assess these inclusion and exclusion criteria. To be conservative, if an exclusion diagnosis or procedure had unspecified laterality, the eye was excluded. The 1-year follow-up analysis cohort eliminated any eye that had any relevant procedure, during follow-up, for which the treatment eye was unspecified.

• **OUTCOMES:** The outcomes for the TVT RCT are listed in Supplemental Table 3.[3] The outcomes for this paper are the TVT IRIS Registry cohort composition, treatment group differences in this cohort, and differences between this cohort and the TVT RCT cohort with respect to baseline characteristics and through 1 year of follow-up. The TVT RCT schedule of follow-up examinations and their "visit windows" are listed in Supplemental Table 3.[3] With only year and week of service, it was impossible to create a 1-day follow-up visit. From each eye's follow-up data in the IRIS Registry, we designated the 1-week, 1-month, 3-month, 6-month, and 1-year follow-up visits using data from the visit that had IOP data closest to the actual follow-up time. If an eye did not have a follow-up visit in a TVT RCT follow-up visit window, that eye was considered to have missed that follow-up visit.

• **STATISTICAL ANALYSIS:** The TVT IRIS Registry treatment groups were compared using independent-samples t tests for continuous variables and χ^2, Fisher exact, or exact χ^2 tests for categorical variables. Comparisons between TVT IRIS Registry and TVT RCT cohorts used χ^2 tests and a 1-sample t test. Statistical analyses were performed

using SAS (SAS Institute, Cary, North Carolina, USA) version 9.4. A P value of ≤.050 was considered statistically significant.

RESULTS

• **BASELINE CHARACTERISTICS:** We received data from the IRIS Registry for 85,416 patients who had a tube and/or a trabeculectomy in 1 or both eyes, from which we identified an initial cohort of 9,404 eyes with both a "qualifying" and "baseline" surgery. We excluded 8,931 (95.0%) eyes to create a baseline cohort of 473 eyes (5.0%) (Figure and Supplemental Table 4; Supplemental Material available at AJO.com). In this baseline cohort, 269 eyes (56.9%) received a tube and 204 eyes (43.1%) received a trabeculectomy, including 21 eyes (4.4%) that had no prebaseline glaucoma medication data. Clinical and demographic characteristics for our baseline cohort (Supplemental Tables 5 and 6; Supplemental Material available at AJO.com) may be compared to tables 3, 4 and 5 in the TVT RCT design and baseline publication.[3] Unlike in the TVT RCT, in our TVT IRIS Registry study, the baseline treatment groups differed significantly on several characteristics (Supplemental Tables 5 and 6).

• **FOLLOW-UP:** Of the 473 eyes in the baseline cohort, 54 lacked data for the 1-year follow-up analyses: 18 eyes had no follow-up IOP data and 36 eyes had a relevant procedure with unspecified laterality during follow-up. There was a significant difference between treatment groups in the TVT IRIS Registry cohort in the percentage of follow-up visits completed (tube 83.8%, trabeculectomy 88.0%, $P = .004$, Table 1). There was also a significantly lower percentage of completed follow-up visits in the TVT IRIS Registry study (85.6%) than in the TVT RCT (96.1%) ($P < .001$, Table 2).[5] Unlike in the TVT RCT,[5] there were no data about patient deaths in the IRIS Registry, so this completeness of follow-up analysis could not adjust for that.

In the TVT IRIS Registry cohort, there were no significant differences between treatment groups in the probability of having data for the 1-year follow-up analyses (tube 95.2%, trabeculectomy 97.6%, $P = .180$) or in the probability of having data for any particular follow-up visit (all $P > .05$) except for the 6-month visit (tube 81.0%, trabeculectomy 91.2%, $P = .002$), but the difference in the probability of having both a 6-month and a 1-year visit was close to significant (tube 75.1%, trabeculectomy 81.9%, $P = .078$, Table 1). There was a significant difference between the TVT-RCT (89.2%)[5] and the TVT IRIS Registry (81.4%) cohorts in the probability of having a 1-year follow-up visit ($P = .011$, Table 2).

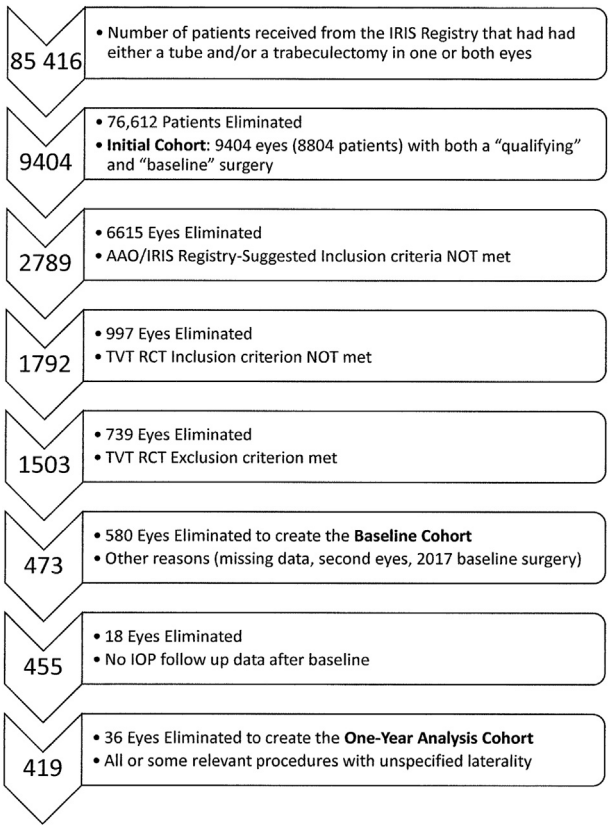

FIGURE. Flowchart of cohort selection for the Tube Versus Trabeculectomy IRIS Registry Study. See Supplemental Table 4 for detail. AAO = American Academy of Ophthalmology; IOP = intraocular pressure; IRIS = Intelligent Research in Sight; RCT = randomized controlled trial; tube = glaucoma drainage implant; TVT = Tube Versus Trabeculectomy.

• **ONE-YEAR ANALYSIS COHORT:** In the TVT IRIS Registry 1-year follow-up analysis cohort of 419 eyes, 236 (56.3%) received a tube and 183 (43.7%) received a trabeculectomy. Baseline clinical and demographic data for this cohort are found in Tables 3 and 4 and Supplemental Table 7 (Supplemental Material available at AJO.com). More trabeculectomy patients were White (61.8% vs 44.9%), whereas more tube patients were Black (35.6% vs 20.8%) or Hispanic (11.0% vs 7.7%) ($P = .003$). More tube patients than trabeculectomy patients had diabetes (41.1% vs 31.2%, $P = .036$), had qualifying surgery that included a trabeculectomy rather than cataract surgery only (42.0% vs 17.5%, $P < .001$), and were using oral carbonic anhydrase inhibitors (35.2% vs 22.4%, $P = .005$). More trabeculectomy eyes than tube eyes had prior laser trabeculoplasty (41.2% vs 28.0%, $P = .003$), had any prior laser procedure (51.4% vs 41.5%, $P = .045$), and were pseudophakic (81.4% vs 72.9%, $P =$

TABLE 1. Treatment Group Follow-up Comparisons in the Tube Versus Trabeculectomy IRIS Registry Study

Description	Trab Group N (%)	Tube Group N (%)	*P* Value
TVT IRIS Registry: Intrastudy comparisons with baseline cohorts			
Follow-up visits completed	898 (88.0%)	1127 (83.8%)	.004**
Had data for the 1-year follow-up analyses	199 (97.6%)	256 (95.2%)	.180
Had a 1-week follow-up visit	177 (86.8%)	230 (85.5%)	.695
Had a 1-month follow-up visit	189 (92.7%)	239 (88.9%)	.163
Had a 3-month follow-up visit	176 (86.3%)	225 (83.6%)	.430
Had a 6-month follow-up visit	186 (91.2%)	218 (81.0%)	.002**
Had a 1-year follow-up visit	170 (83.3%)	215 (79.9%)	.346
Had both 6-month & 1-year follow-up visits	167 (81.9%)	202 (75.1%)	.078

IRIS = Intelligent Research in Sight; trab = trabeculectomy; tube = glaucoma drainage implant; TVT = Tube Versus Trabeculectomy.
All χ^2 tests.
*$P \leq .05$, ** $P \leq .01$, *** $P \leq .001$.

TABLE 2. Interstudy Follow-up Comparisons in the Tube Versus Trabeculectomy IRIS Registry Study

Group Description	TVT IRIS Registry N (%)		Alternate Study N (%)		*P* Value
Interstudy comparisons with the baseline cohorts: categorical variables					
Follow-up visits completed	All Eyes	2025 (85.6%)	TVT RCT: All Eyes	1222 (96.1%)	<.001***
Had a 1-year follow-up visit	All Eyes	385 (81.4%)	TVT RCT: All Eyes	189 (89.2%)	.011*
No medication at baseline	All Eyes	21 (4.4%)	TVT RCT: All Eyes	3 (1.4%)	.047*
No medication at baseline	All Eyes	21 (4.4%)	ABC RCT: All Eyes	4 (1.5%)	.028*
Stratum: previous cataract extraction only	All Eyes	331 (70.0%)	TVT RCT: All Eyes	94 (44.3%)	<.001***
Stratum: previous cataract extraction only	Trabs	169 (82.8%)	TVT RCT: Trabs	48 (45.7%)	<.001***
Stratum: previous cataract extraction only	Tubes	162 (60.2%)	TVT RCT: Tubes	41 (38.3%)	<.001***
	TVT IRIS Registry mean (SD)		Alternate study mean (SD)		
Interstudy comparisons with the baseline cohorts: continuous variable					
Previous incisional surgery: total months before baseline	All eyes	12 (9)	TVT RCT: All Eyes	57 (52)	<0.001***[a]

ABC = Ahmed Baerveldt Comparison Study; IRIS = Intelligent Research in Sight; RCT = randomized controlled trial; SD = standard deviation; trab = trabeculectomy; tube = glaucoma drainage implant; TVT = Tube Versus Trabeculectomy.
All χ^2 tests except as noted.
*$P \leq .05$, ** $P \leq .01$, *** $P \leq .001$.
[a]One-sample *t* test.

.041). Tube eyes had a higher mean number of previous incisional surgeries (1.9 vs 1.5, $P < .001$), mean logMAR VA (0.67 vs 0.47, $P < .001$), and more severe glaucoma ($P = .035$) than trabeculectomy eyes. In trabeculectomy eyes, the previous incisional surgery was a mean of 1.9 months closer to the baseline surgery than in tube eyes (13.4 months vs 11.5 months, $P = .033$). Also, the mean time from the previous incisional surgery to the baseline surgery in the TVT RCT cohort was 57 months,[3] whereas in the TVT IRIS Registry cohort this was only 12 months ($P < .001$, Table 2). Treatment group differences for several other variables were close to significant (eg, age and IOP; Tables 3 and 4, and Supplemental Table 7).

TABLE 3. Baseline Demographic and Clinical Characteristics of the 1-Year Follow-up Analysis Cohort in the Tube Versus Trabeculectomy IRIS Registry Study: Categorical Variables

Baseline Demographic and Clinical Variables	Values	Tube Group N = 236 (56.3%)	Trab Group N = 183, (43.7%)	*P* Value
Sex	Male	113 (47.9%)	84 (45.9%)	.687
Race	White	106 (44.9%)	113 (61.8%)	.003**
	Black	84 (35.6%)	38 (20.8%)	
	Hispanic	26 (11.0%)	14 (7.7%)	
	Asian	6 (2.5%)	3 (1.6%)	
	Other/Mixed	14 (5.9%)	15 (8.2%)	
Patient had hypertension	Yes	172 (72.9%)	119 (65.0%)	.083
Patient had diabetes	Yes	97 (41.1%)	57 (31.2%)	.036*
Laterality	Right	114 (48.3%)	91 (49.7%)	.773
Previous laser any	Yes	98 (41.5%)	94 (51.4%)	.045*
Previous LTP	Yes	66 (28.0%)	77 (42.1%)	.003**
Previous LPI	Yes	12 (5.1%)	9 (4.9%)	.938
Previous laser other	Yes	28 (11.9%)	23 (12.6%)	.827
Diagnosis group	POAG	204 (86.4%)	159 (86.9%)	.654[b]
	CACG	12 (5.1%)	8 (4.4%)	
	PXF	10 (4.2%)	12 (6.6%)	
	Pigmentary	4 (1.7%)	2 (1.1%)	
	Other glaucoma	6 (2.5%)	2 (1.1%)	
Lens status	Phakic	64 (27.1%)	34 (18.6%)	.041*
	Pseudophakic	172 (72.9%)	149 (81.4%)	
Diplopia	Yes	4 (1.7%)	2 (1.1%)	.700[a]
Stratum	Previous cataract extraction	137 (58.1%)	151 (82.5%)	<.001***
	Previous trab or combined procedure	99 (42.0%)	32 (17.5%)	
Glaucoma medications any	Yes	227 (96.2%)	175 (95.6%)	.774
Prostaglandin analogue	Yes	191 (80.9%)	154 (84.2%)	.391
Beta blocker	Yes	198 (83.9%)	148 (80.9%)	.418
Topical CAI	Yes	187 (79.2%)	145 (79.2%)	1.000
Alpha agonist	Yes	184 (78.0%)	135 (73.8%)	.318
Oral CAI	Yes	83 (35.2%)	41 (22.4%)	.005**
Cholinergic agonist	Yes	0 (0.0%)	1 (0.6%)	.437[a]
Other glaucoma medication	Yes	2 (0.9%)	1 (0.6%)	1.000[a]
Glaucoma medication unknown	Yes	13 (5.5%)	13 (7.1%)	.502
Glaucoma stage (99 eyes did not have baseline glaucoma stage data)	Mild	12 (6.6%)	20 (14.6%)	.035*
	Moderate	43 (23.5%)	36 (26.3%)	
	Severe	128 (70.0%)	81 (59.1%)	

CACG = chronic angle closure glaucoma; CAI = carbonic anhydrase inhibitors; IRIS = Intelligent Research in Sight; LPI = laser iridotomy/iridectomy; LTP = laser trabeculoplasty; med =medication; n = number; POAG = Primary open angle glaucoma; PXF = pseudoexfoliation glaucoma; trab = trabeculectomy; tube = glaucoma drainage implant.

All χ^2 tests except as noted.

*$P \leq .05$, ** $P \leq .01$, *** $P \leq .001$.

[a]Fisher exact test.

[b]Exact χ^2 test.

DISCUSSION

THE COMPARISON OF OUTCOMES BY TREATMENT GROUP may be confounded by attrition bias, caused by differences in follow-up, since there was a significant difference between the TVT IRIS Registry treatment groups in the percentage of follow-up visits completed (Table 1). Attrition bias "occurs when the duration of follow-up differs systematically between the compared treatment groups," but it can also occur when there is a difference in the overall pattern of follow-up visits completed.[7] Comparisons of the TVT RCT and TVT IRIS Registry studies also indicated significant differences, both in the percentage of follow-up visits completed and in the probability of an eye having a 1-year follow-up visit (Table 2). The percentage of missed follow-up visits in the TVT IRIS Registry

TABLE 4. Baseline Demographic and Clinical Characteristics of the 1-Year Follow-up Analyses Cohort in the Tube Versus Trabeculectomy IRIS Registry Study: Continuous Variables

Compare TVT IRIS Registry Treatment Groups (Tube N = 236, 56.3%; Trab N = 183, 43.7%)

Baseline Demographic and Clinical Variables	Group	Mean (Standard Deviation)	5-Number Summary	Mean Diff. (95% CI Diff.)	*P* Value
Age at baseline surgery	Tube	70.0 (9.8)	38, 64, 71, 77, 85	-1.8	.056
	Trab	71.8 (8.8)	40, 67, 73, 78, 85	(-3.6, 0.0)	
Intraocular pressure (mm Hg)	Tube	26.6 (6.5)	18, 21, 25, 32, 40	1.2	.054
	Trab	25.3 (6.4)	18, 20, 23, 30, 40	(0.0, 2.5)	
Total glaucoma medications	Tube	3.6 (1.3)	0, 3, 4, 4, 6	0.2	.170
	Trab	3.4 (1.2)	0, 3, 4, 4, 5	(-0.1, 0.4)	
Total number of previous incisional surgeries	Tube	1.9 (1.1)	1, 1, 2, 2, 10	0.4	<.001***
	Trab	1.5 (0.7)	1, 1, 1, 2, 4	(0.2, 0.5)	
Most recent incisional surgery: number months before baseline	Tube	11.5 (10.2)	0, 4.5, 9, 16.4, 78.2	-1.9	.033*
	Trab	13.4 (8.4)	0, 7.1, 11.5, 19.4, 51.5	(-3.7, -0.2)	
Approximate ETDRS letters	Tube	51.8 (28.9)	-50, 40, 61, 70, 90	-10.0	<.001***
	Trab	61.7 (23.8)	-30, 55, 67, 76, 90	(-15.0, -4.9)	
LogMAR visual acuity	Tube	0.67 (0.58)	-0.1, 0.3, 0.48, 0.91, 2.7	0.2	<.001***
	Trab	0.47 (0.48)	-0.1, 0.18, 0.36, 0.6, 2.3	(0.1, 0.3)	

5-number summary = minimum, 25th percentile, median, 75th percentile, maximum; CI = confidence interval; Diff. = difference; ETDRS = Early Treatment for Diabetic Retinopathy Study; IRIS = Intelligent Research in Sight; LogMAR = logarithm of the minimum angle of resolution; trab = trabeculectomy; tube = glaucoma drainage implant; TVT = Tube Versus Trabeculectomy.

All independent-samples *t* tests.

*$P \leq .05$, ** $P \leq .01$, *** $P \leq .001$.

study (14.4%) was, in fact, greater than the percentage of missed follow-up visits during the entire 5 years of the TVT RCT (13.0%).[8] From a patient care perspective, the TVT IRIS Registry "missed visits" are not necessarily a reason for concern, since there are no expectations that patient follow-up schedules in clinical practice would coincide with RCT follow-up protocols. However, for clinical studies that seek to evaluate patient outcomes at specific time points, the relatively larger percentage of IRIS Registry patients without data for specific TVT RCT follow-up visits indicates that using EHR data might result in more patients without timely visits for follow-up analyses. This may increase bias in EHR-based treatment evaluations compared to RCT-based treatment evaluations. Although alternate statistical analysis approaches, such as using survival analysis to account for variable follow-up time, and/or increasing the acceptable "window" for a follow-up visit, might address some clinical questions, these are likely to be more useful for long-term studies than short-term studies. For example, using a 5.5-year visit for a 5-year outcome is likely to be more acceptable than using a 9-month visit for a 3-month outcome, although each visit is 6 months "late."

IOP readings in the TVT RCT were obtained using Goldmann applanation tonometry, considered by providers as the gold standard to measure IOP,[9] and IOP readings were repeated until 2 measurements (subsequently averaged) differed by 1 mm Hg or less.[3] In the IRIS Registry there was no indication of how IOP was measured. In a recent survey, Junk and associates[9] found that most glaucoma specialists preferred Goldmann applanation tonometry (98%) and considered this the most accurate method (82%) to measure IOP. However, Goldmann applanation tonometry was seldom used by technicians (2%), who most frequently measured IOP with the air-puff (42%) or iCare Rebound (31%) tonometer, which most glaucoma specialists did not consider as accurate as Goldmann applanation tonometry.[9] Since a majority of glaucoma specialists reported that technicians perform some IOP measurements (82%),[9] the IOP data from the IRIS Registry is likely to be less accurate than the IOP data from the TVT RCT. It probably will also be less precise, because usually the IRIS Registry contained only a single IOP measurement, whereas RCTs often average at least 2 IOP measurements, and mandate that these differ by 1 mm Hg or less.[3,10,11] In general, EHR data may lack the accuracy and precision of RCT-generated data.

Parkinson and associates[12] noted "that changing treatment practices can limit the ability to interpret differences in health outcomes" using "real-world observational data." Both the 1-year (published in 2007)[5] and 5-year (published in 2012)[8] TVT RCT results indicated superior outcomes for eyes that received a tube, although the 3-year results were equivocal.[13] It is impossible to know how evidence

from the TVT RCT and other subsequent research, such as the Ahmed Baerveldt Comparison Study (ABC) RCT and the Ahmed Versus Baerveldt Study (AVB) RCT, affected treatment choice for eyes in the TVT IRIS Registry cohort.[14,15] However, there is evidence that the choice of a tube or a trabeculectomy has changed over time. In surveys about practice preferences, for eyes that had a prior trabeculectomy, in 2008 American Glaucoma Society (AGS) members indicated that 49% would perform a second trabeculectomy and 46% would insert a tube, whereas in 2016 only 20% indicated that they would perform a second trabeculectomy and 71% would insert a tube. For eyes that had prior phacoemulsification, in 2008 AGS members indicated that 74% would perform a trabeculectomy and 15% would insert a tube, whereas in 2016 only 60% indicated that they would perform a trabeculectomy and 34% would insert a tube.[16] For eyes with either prior (qualifying) procedure, in 2016 compared to 2008, fewer AGS members would choose a trabeculectomy and more would choose a tube.

In the TVT RCT at baseline, there were no significant treatment group differences,[3] whereas in our TVT IRIS Registry cohort, the treatment groups differed significantly on several characteristics in both the baseline and 1-year analysis cohorts (Tables 3 and 4 and Supplemental Tables 5–7). Selection bias, owing to nonrandom treatment assignment, is one obvious explanation. In our 1-year analysis cohort, more trabeculectomy patients were White, whereas more tube patients were Black or Hispanic ($P = .003$). When the initial glaucoma surgery was a trabeculectomy, the Advanced Glaucoma Intervention Study found a higher risk of failure among Black patients than White patients.[17] A recent literature review found that "African descent is a well-documented risk factor for glaucoma development, progression, and medical and surgical failure," with reports of "decreased surgical success in Black patients" after both trabeculectomy and tube-shunt surgery.[18] With Baerveldt tubes, there was little evidence of an increased risk of failure among Black patients compared to others, but with Ahmed tubes, an increased risk of failure among Black patients was found.[18] The TVT RCT, which found reduced risk of failure in tubes compared to trabeculectomy at 1 year, used only Baerveldt 350 tubes and had no significant treatment group differences by race ($P = .53$).[5] Analysts at the Academy searched the IRIS Registry's free text fields for an indication of tube type, but found few, so the TVT IRIS Registry cohort included all tube types. The comparison of outcomes by race of the TVT IRIS Registry and the TVT RCT studies might be confounded because, in the IRIS Registry data, Black patients, for whom there exists evidence of greater risk of surgical failure,[18] were significantly more likely to receive a tube than White patients, and these tubes would have included any tube type billed with CPT codes 66179 or 66180, such as Ahmed, Molteno, and Baerveldt 250 or 350 tubes.

In our baseline cohort, 21 (4.4%) eyes had no prebaseline glaucoma medication data, whereas in the TVT RCT, only 3 (1.4%) patients had no prebaseline glaucoma medications,[3] and this difference was statistically significant ($P = .047$). In the ABC RCT, only 4 (1.5%) patients had no prebaseline glaucoma medications,[10] which was also significantly different from our cohort ($P = .028$) (Table 2). All TVT RCT patients and 80% of the ABC patients had previous intraocular surgery, and both studies included only patients for whom additional medical therapy was not possible.[3,10] All TVT IRIS Registry study patients had previous intraocular surgery, but there was no indication, in the IRIS Registry, of why the baseline surgery was performed, although incisional surgery is often preformed when additional medical therapy is not possible. Therefore, although 21 eyes without prebaseline glaucoma medication data is a small percentage (4.4%) of the TVT IRIS Registry cohort, some TVT IRIS Registry study eyes may not have been eligible for the TVT RCT.

The CPT procedure codes for trabeculectomy did not specify whether an antifibrotic agent was used, so it was not possible to categorize the TVT IRIS Registry eyes into the 4 TVT RCT strata (Supplemental Table 3).[3] Instead, we classified eyes into 2 strata: previous cataract extraction only and previous trabeculectomy or combined procedure (with or without an antifibrotic agent). In the TVT RCT, 23% of eyes had a qualifying surgery of a previous trabeculectomy or combined procedure without an antifibrotic agent,[3] whereas, because trabeculectomy with an antifibrotic agent is the current standard of care,[18] it is likely that few or no eyes in the TVT IRIS Registry study would have been in this stratum. Also, only 5 years of data were available in the IRIS Registry database, so when the qualifying surgery included a (failed) previous trabeculectomy, the TVT IRIS Registry cohort might include disproportionately more eyes at a high risk for earlier failure for any bleb-forming procedure (either a tube or a trabeculectomy) than the TVT RCT cohort.

There is evidence that it is difficult to select an RWD cohort similar to that from an RCT,[1,2] which is clearly the first requirement in replicating the results of an RCT with RWD. We applied the TVT RCT inclusion criteria, including that for qualifying surgery,[3] to the IRIS Registry data, but the TVT IRIS Registry cohort still differed significantly from the TVT RCT cohort with respect to the types of qualifying surgery. In our baseline cohort, 331 (70.0%, Supplemental Table 5) eyes had a qualifying surgery of previous cataract extraction only, whereas in the TVT RCT, significantly fewer (94, 44.3%) eyes had this qualifying surgery ($P < .001$, Table 2). Within the TVT IRIS Registry study, the qualifying surgeries differed as well, with 82.8% of the trabeculectomy eyes having previous cataract extraction only, whereas only 60.2% of the tube eyes had this qualifying surgery ($P < .001$, Supplemental Table 5), and both of these percentages were significantly different from the TVT RCT percentages for the analogous treatment groups (both $P < .001$, Table 2). It is possible that in clinical practice, a previous trabeculectomy failure makes a subsequent glaucoma surgery choice less likely to

be another trabeculectomy. For example, in the Primary TVT RCT 3-year results, 10 eyes in the trabeculectomy group required a reoperation for glaucoma, but only 1 received a second trabeculectomy.[19] Therefore it is possible that, in the IRIS Registry cohort, the eyes that received a tube had more advanced glaucoma than the eyes that received a trabeculectomy, since nearly 40% of eyes that received a tube had a previous glaucoma surgery (trabeculectomy) whereas fewer than 20% of eyes that received a trabeculectomy had a previous glaucoma surgery. In addition, in the TVT IRIS Registry 1-year analysis cohort, tube eyes had a higher mean number of previous incisional surgeries (1.9 vs 1.5, $P < .001$), had higher mean logMAR VA (0.67 vs 0.47, $P < .001$), had more severe glaucoma ($P = .035$) than trabeculectomy eyes, and were more likely to belong to patients with diabetes (41.1% vs 31.2%, $P = .036$) than trabeculectomy eyes (Tables 3 and 4).

In summary, the TVT IRIS Registry study faced several limitations in creating a cohort of eyes analogous to that in the TVT RCT. The EHR-based IRIS Registry data seldom had an indication of which type of tube was used, and the CPT procedure codes for trabeculectomy did not indicate whether an antifibrotic agent was used. Certain TVT RCT inclusion and exclusion criteria could not be ascertained from the IRIS Registry mainly because there were no relevant or specific ICD or CPT codes (Supplemental Table 2), a problem identified for potential replication of many RCTs with EHR data.[1] Even if all inclusion and exclusion criteria could have been applied, the resulting RWD-cohort still might have been dissimilar and unable to replicate the TVT RCT cohort's results.[2] As a result, there were many significant differences between the tube and trabeculectomy eyes in the TVT IRIS Registry cohort and between that cohort and the TVT RCT cohort. Other data used in the TVT RCT were also not available in the IRIS Registry, including whether a patient died during follow-up or was pregnant or nursing at baseline, transient complications (eg, bleb leaks), visual field,[20] eye motility,[21] and quality of life.[22]

Routine replication of RCT results with RWD may prove to be difficult or impossible, but fortunately, this is not the primary goal of RWD research. "Can RWD replicate an RCT?" is an academic question, not a clinical question. The inability to replicate RCT cohorts and results should not imply that RWD will not be useful for comparing treatments and informing clinical practice. Results can differ without either being invalid. Different results from distinct cohorts are attributable, in part, to the heterogeneity of treatment effect,[23] which exists because clinical and demographic characteristics, other than treatment choice, also influence patient outcomes. These heterogeneity of treatment effects may be one reason why treatment comparisons with RWD would provide critical information about real-world effectiveness to complement information about treatment efficacy provided by RCTs.

The inclusion and exclusion criteria of an RCT often create a research cohort that does not represent the entire relevant patient population. Exclusion of some patients from an RCT occurs to reduce confounding, to increase internal validity, and to isolate the treatment effect to compare efficacy.[24] This, however, limits the generalizability (external validity) of the RCT to all patients for whom the treatments might be considered, which is a problem that has long been noted.[2,24] RCT results often cannot be generalized to the most complex and unusual patients,[24] for whom clinicians have the least experience and are, therefore, in the greatest need of guidance.

A complete database of RWD for ophthalmology in the United States, which is the goal of the IRIS Registry, will contain the clinical experience of the patient population for existing ophthalmology practice patterns. Data will be available for many more patients and for longer follow-up time than is feasible for RCTs,[25] without the researcher needing to expend the resources to collect it. There would be fewer concerns about the generalizability of results from research that included the entire population of treated patients. Also, RCTs often use a surrogate outcome,[25] such as IOP reduction, because this is expected to change in the timeframe of the RCT and it is a known risk factor for glaucoma progression. Extensive follow-up would enable tracking patients for years, even with different providers, and would allow assessment of outcomes that change more slowly and less consistently, but which are more patient-centered, such as VA. Big-RWD will enhance the assessment of complications and permit treatment comparisons for specific subgroups of patients, neither of which most RCTs are powered to assess. In addition to informing clinical practice, RWD will be useful for informing clinical research to generate and pretest hypotheses and to assess how practice patterns change based on previous research results.[24]

Future research by these authors will include comparing clinical outcomes between the tube and trabeculectomy eyes in the TVT IRIS Registry cohort and comparing this cohort's clinical outcomes with those from the TVT RCT cohort. Additional future research could include creating a larger TVT IRIS Registry cohort using additional years of data, since a larger sample might enable the use of propensity scores and/or multivariate adjustment for unbalanced baseline variables. One blessing of the IRIS Registry is that the most significant resource needed for a future project might be patience, since the programming exists, and additional data are collected continuously. If possible, a future IRIS Registry cohort could include only patients who had their baseline surgery at a large clinical center, such as those that participated in the TVT RCT, to create a TVT IRIS Registry cohort from a population that was more similar to that included in the TVT RCT. As with the TVT RCT, comparisons between 3- and 5-year outcomes could also be performed with updated data.

In conclusion, using IRIS Registry data, we created a cohort of eyes as similar as possible to those included in the TVT RCT. Owing to the retrospective nature of the IRIS Registry data and the lack of random treatment

assignment, there were several significant differences in this cohort between eyes that received tubes and trabeculectomies, whereas there had been no significant treatment group differences at baseline in the TVT RCT. There were also significant differences between the eyes in our cohort and those in the TVT RCT cohort with respect to baseline characteristics and follow-up. It remains to be seen how these differences might affect clinical outcomes in the TVT IRIS Registry cohort and the comparison of these outcomes with those of the TVT RCT.

FUNDING/SUPPORT: THIS WORK WAS SUPPORTED BY A 2018 RESEARCH TO PREVENT BLINDNESS (NEW YORK, NEW YORK, USA)/ American Academy of Ophthalmology (San Francisco, California, USA) Award for IRIS Registry Research, the National Institutes of Health (Core Grant number P30EY014801), and a Research to Prevent Blindness (New York, New York, USA) Unrestricted Grant. The sponsor or funding organization participated in the design of the study, conducting the study, data collection, data management, and review and approval of the manuscript. Dr Lum is an employee of the American Academy of Ophthalmology. Dr Kelly was an employee of the American Academy of Ophthalmology during part of the time in which this research was conducted. Financial Disclosures: During the conduct of the study: (1) Drs Vanner, Sun, Persad, Parrish, Chang, and Gedde and Mr McSoley and Mr Feuer report grant support from the National Institutes of Health (grant number P30EY014801) and an Unrestricted Grant from Research to Prevent Blindness, New York, New York, USA; (2) Dr Vanner reports a 2018 Research to Prevent Blindness (New York, New York, USA)/ American Academy of Ophthalmology (San Francisco, California, USA) Award for IRIS Registry Research; (3) Drs Chang and Vanner report a grant from the American Glaucoma Society, San Francisco, CA 2018 IRIS Registry-AGS Research Initiative; (4) Mr Feuer and Dr Vanner report grant support from the National Eye Institute (grant number NEI UG1 EY024247); (5) Dr Parrish and Mr Feuer report grant support from the National Eye Institute (grant number NEI 1UG1 EY025183-01); (6) Mr Feuer reports grant support from the National Eye Institute (grant number NEI R01 EY019077); and (7) Dr Sun reports grant support from the National Eye Institute (grant number K12EY031372). All authors attest that they meet the current ICMJE criteria for authorship.

REFERENCES

1. Bartlett VL, Dhruva SS, Shah ND, Ryan P, Ross JS. Feasibility of using real-world data to replicate clinical trial evidence. *JAMA Netw Open* 2019;2(10):e1912869.
2. Averitt AJ, Weng C, Ryan P, Perotte A. Translating evidence into practice: eligibility criteria fail to eliminate clinically significant differences between real-world and study populations. *NPJ Digit Med* 2020;67:1–10.
3. Gedde SJ, Schiffman JC, Feuer WJ, Parrish RK 2nd, Heuer DK, Brandt JD, for Tube Versus Trabeculectomy Study Group. The tube versus trabeculectomy study: design and baseline characteristics of study patients. *Am J Ophthalmol* 2005;140(2):275–287.
4. Chiang MF, Sommer A, Rich WL, Lum F, Parke DW 2nd. The 2016 American Academy of Ophthalmology IRIS® Registry (Intelligent Research in Sight) database: characteristics and methods. *Ophthalmology* 2018;125(8):1143–1148.
5. Gedde SJ, Schiffman JC, Feuer WJ, Herndon LW, Brandt JD, Budenz DL, for Tube Versus Trabeculectomy Study Group. Treatment outcomes in the Tube Versus Trabeculectomy Study after one year of follow-up. *Am J Ophthalmol* 2007; 143(1):9–22.
6. Gedde SJ, Herndon LW, Brandt JD, Budenz DL, Feuer WJ, Schiffman JC, for Tube Versus Trabeculectomy Study Group. Surgical complications in the Tube Versus Trabeculectomy Study during the first year of follow-up. *Am J Ophthalmol* 2007;143(1):23–31.
7. Tur C, Kalincik T, Oh J, et al. Head-to-head drug comparisons in multiple sclerosis: urgent action needed. *Neurology* 2019;93(18):793–809.
8. Gedde SJ, Schiffman JC, Feuer WJ, Herndon LW, Brandt JD, Budenz DL, for Tube Versus Trabeculectomy Study Group. Treatment outcomes in the Tube Versus Trabeculectomy (TVT) study after five years of follow-up. *Am J Ophthalmol* 2012;153(5):789–803.e782.
9. Junk AK, Chang TC, Vanner EA, Chen T. Current trends in tonometry and tonometer tip disinfection. *J Glaucoma* 2020; 29(7):507–512.
10. Barton K, Gedde SJ, Budenz DL, Feuer WJ, Schiffman J, for Ahmed Baerveldt Comparison Study Group. The Ahmed Baerveldt Comparison Study methodology, baseline patient characteristics, and intraoperative complications. *Ophthalmology* 2011;118(3):435–442.
11. Gedde SJ, Chen PP, Heuer DK, et al; for Primary Tube Versus Trabeculectomy Study Group. The Primary Tube Versus Trabeculectomy Study: methodology of a multicenter randomized clinical trial comparing tube shunt surgery and trabeculectomy with mitomycin C. *Ophthalmology* 2018; 125(5):774–781.
12. Parkinson B, Viney R, Haas M, Goodall S, Srasuebkul P, Pearson SA. Real-world evidence: a comparison of the Australian herceptin program and clinical trials of trastuzumab for HER2-positive metastatic breast cancer. *Pharmacoeconomics* 2016;34(10):1039–1050.
13. Gedde SJ, Schiffman JC, Feuer WJ, Herndon LW, Brandt JD, Budenz DL, for Tube Versus Trabeculectomy Study Group. Three-year follow-up of the tube versus trabeculectomy study. *Am J Ophthalmol* 2009;148(5):670–684.
14. Christakis PG, Kalenak JW, Zurakowski D, et al. The Ahmed Versus Baerveldt study: one-year treatment outcomes. *Ophthalmology* 2011;118(11):2180–2189.
15. Budenz DL, Barton K, Feuer WJ, et al; for Ahmed Baerveldt Comparison Study Group. Treatment outcomes in the Ahmed Baerveldt Comparison Study after 1 year of follow-up. *Ophthalmology* 2011;118(3):443–452.
16. Vinod K, Gedde SJ, Feuer WJ, et al. Practice preferences for glaucoma surgery: a survey of the American Glaucoma Society. *J Glaucoma* 2017;26(8):687–693.
17. The AGIS Investigators. The Advanced Glaucoma Intervention Study (AGIS): 9. Comparison of glaucoma outcomes in black and white patients within treatment groups. *Am J Ophthalmol* 2001;132(3):311–320.
18. Taubenslag KJ, Kammer JA. Outcomes disparities between black and white populations in the surgical management of glaucoma. *Semin Ophthalmol* 2016;31(4):385–393.
19. Gedde SJ, Feuer WJ, Lim KS, et al; for Primary Tube Versus Trabeculectomy Study Group. Treatment

outcomes in the Primary Tube Versus Trabeculectomy Study after 3 years of follow-up. *Ophthalmology* 2020; 127(3):333–345.
20. Swaminathan SS, Jammal AA, Kornmann HL, et al; for Tube Versus Trabeculectomy Study Group. Visual Field Outcomes in the Tube Versus Trabeculectomy Study. *Ophthalmology* 2020;127(9):1162–1169.
21. Rauscher FM, Gedde SJ, Schiffman JC, Feuer WJ, Barton K, Lee RK, for Tube Versus Trabeculectomy Study Group. Motility disturbances in the tube versus trabeculectomy study during the first year of follow-up. *Am J Ophthalmol* 2009; 147(3):458–466.
22. Kotecha A, Feuer WJ, Barton K, Gedde SJ, for Tube Versus Trabeculectomy Study Group. Quality of life in the Tube Versus Trabeculectomy Study. *Am J Ophthalmol* 2017;176:228–235.
23. Kent DM, Nelson J, Dahabreh IJ, Rothwell PM, Altman DG, Hayward RA. Risk and treatment effect heterogeneity: re-analysis of individual participant data from 32 large clinical trials. *Int J Epidemiol* 2016;45:2075–2088.
24. Sherman SE, Anderson SA, Dal Pan GJ, et al. Real-world evidence — what is it and what can it tell us? *N Engl J Med* 2016; 375(23):2293–2297.
25. Frieden TR. Evidence for health decision making —beyond randomized, controlled trials. *N Engl J Med* 2017;377(5):465–475.