NIH Collaboratory

Health Care Systems Research Collaboratory

"Table 1"

Slides by
Rachel Richesson,
co-Chair, Phenotype, Data Standards, and Data Quality Core

Table 1. Participant Characteristics at Baseline

	D	Placebo (n = 73)	*P* Value
		63.0 (9.3)	.41
		40 (54)	.13
		63 (86)	.07
		2 (3)	
		8 (11)	
		0 (0)	

Table 1. Baseline Demographics and Clinical Characteristics

Characteristics	Participants by Study Group, No. (%) Nifedipine (n = 201)	Placebo (n = 205)	Nonrandomization (n = 230)
Age, mean (SD), y	30.2 (5.1)	30.2 (5.1)	28.7 (5.5)[a]
Body mass index, mean (SD)[b]	23.3 (4.7)	23.3 (4.3)	23.5 (5.0)
White race	166 (82.6)	155 (75.6)	166 (72.2)
Nulliparous	116 (57.7)		
Prior preterm birth	39 (19.4)		
Gestational age at study entry, mean (SD), wk	29.2 (1.7)		
Multiple gestation	42 (20.9)		
Twins	40 (19.9)		
Triplets	2 (1.0)		
pPROM at study entry	53 (26.4)		
Vaginal bleeding at study entry	38 (18.9)		
Additional tocolysis	12 (6.0)		
Vaginal examination at study entry	(n = 134)		
Dilatation, median (IQR), cm	0 (0-1)		
Cervical length, median (IQR), mm	25 (15-35)		

Abbreviations: IQR, interquartile range; NA, not applicable; pPROM, preterm premature rupture of membranes; SD, standard deviation.
[a] $P = .01$.
[b] Body mass index is calculated as weight in kilograms divided by height in meters squared.

Table 1. Patient Characteristics

Characteristic	No. (%) Paracetamol (n = 33)	Morphine (n = 38)	*P* Value
Sex			
Male	18 (54.5)	26 (68.4)	.23
Female	15 (45.5)	12 (31.6)	
Age at surgery, median (IQR), d	5 (1.5-64.5)	20 (1.8-87.5)	.50
Age, d			
≤10	17 (51.5)	18 (47.4)	.73
>10	16 (48.5)	20 (52.6)	
Weight, mean (SD), kg	3.8 (1.3)	4.4 (2.0)	.17
Duration of surgery, mean (SD), min	172.1 (83.7)	156.6 (87.9)	.45
Surgical procedure			
Thoracic	5 (15.2)	11 (28.9)	.17
Abdominal	28 (84.8)	27 (71.1)	
Postoperative mechanical ventilation	15 (45.5)	14 (36.8)	.46
Duration of postoperative ventilation, median (IQR), h	34 (15-45)	23 (16-45)	.43
Surgical stress score, median (IQR)	10 (9-11)	10 (9-11)	.75
PRISM3, median (IQR)	2 (0-4.5)	3.0 (0-5.0)	.91
PIM2, median (IQR), % risk of mortality	1.3 (0.6-1.9)	1.4 (0.7-2.4)	.34

Abbreviations: IQR, interquartile range; PIM2, Pediatric Index of Mortality 2; PRISM3, Pediatric Risk of Mortality 3.

Table 1. Patient D

C		
Patient demographic		
Age, median (IQR		
White race		
Weight, median (		
Body mass index		
Body mass index		
Male sex		
Medical history		
History of hyperte		
History of diabete		
Current or histor		
Current smok		
History of chronic		
History of periphe		
Previous median sternotomy	52 (6.9)	42 (5.6)
History of TIA or stroke	77 (10.2)	81 (10.8)
History of myocardial infarction	233 (31.0)	245 (32.7)
History of congestive heart failure	89 (11.8)	90 (12.0)
History of hyperlipidemia	619 (82.2)	607 (81.0)
Steroid use ≤1 mo prior to surgery	28 (3.7)	33 (4.4)
Receiving dialysis preoperatively	4 (0.5)	2 (0.3)
Preoperative diagnostic values		
Left ventricular ejection fraction, median (IQR), %	55 (45-60)	55 (45-60)
Serum glucose, median (IQR), mg/dL	125 (101-160)	124 (103-167)
Serum hemoglobin A_{1c}, median (IQR), %	6.5 (5.9-7.6)	6.6 (5.9-7.7)
Hematocrit, median (IQR), %	39 (36-42)	39 (36-42)
Serum creatinine, median (IQR), mg/dL	1.0 (0.9-1.3)	1.0 (0.9-1.2)
Preoperative core temperature, median (IQR), °C	97.6 (97.0-98.2)	97.7 (97.0-98.2)
Preoperative hospital stay, median (IQR), d	1.0 (0-3.0)	1.0 (0-3.0)
Parsonnet risk score, median (IQR)[b]	9.0 (6.0-14.5)	9.0 (6.0-16.0)

Abbreviations: IQR, interquartile range; TIA, transient ischemic attack.
SI conversion factors: To convert creatinine to µmol/L, multiply by 88.4; glucose to mmol/L, multiply by 0.0555.
[a] Unless otherwise indicated.
[b] Theoretical range is 0 to 148; 50% in Parsonnet et al[11] had a score between 0 and 9.

ume, mean (SD), mm³	4212 (1349)	4740 (1273)	.03
ckness, mean (SD), mm²	1.8 (0.4)	1.8 (0.3)	.28
olume, cm³			
an (SD)	8.5 (15.0)	9.6 (14.0)	.68
dian (IQR)	1.3 (0.0-10.7)	2.3 (0.0-15.2)	.75
-hydroxyvitamin D, mean g/mL	22.7 (11.4)	21.9 (8.3)	.62
e width, mean (SD), mm	5.0 (1.8)	5.1 (1.7)	.66

s: BMD, bone mineral density; BMI, body mass index, calculated as weight in kilograms divided by height squared; BML, bone marrow lesion; IQR, interquartile range; KL, Kellgren-Lawrence; WOMAC, Western d McMaster Universities.
[a] Knee-specific pain ranges from 0 to 20 with 0 indicating no pain and knee-specific function ranges from 0 to 68, with 0 indicating no pain with activity.

Project Aim

- To identify important person characteristics and clinical features, along with explicit definitions and representations, for the reporting of baseline characteristics of research populations in interventional and observational studies.

Table 1 supports:

- Submission of data sets and data results from NIH-funded studies for archival and secondary use and for analyses and comparisons across trials
- Standardizing reporting of results from NIH-funded studies to ClinicalTrials.gov
- Better practices for describing populations in submissions to medical journals
- Conduct of multisite pragmatic trials
- Conduct of multisite observational studies
- Others?

Question 1:
What characteristics should be on Table 1?

Question 2:
How should they be defined? represented?

Table 1. Patient Demographics and Baseline Characteristics

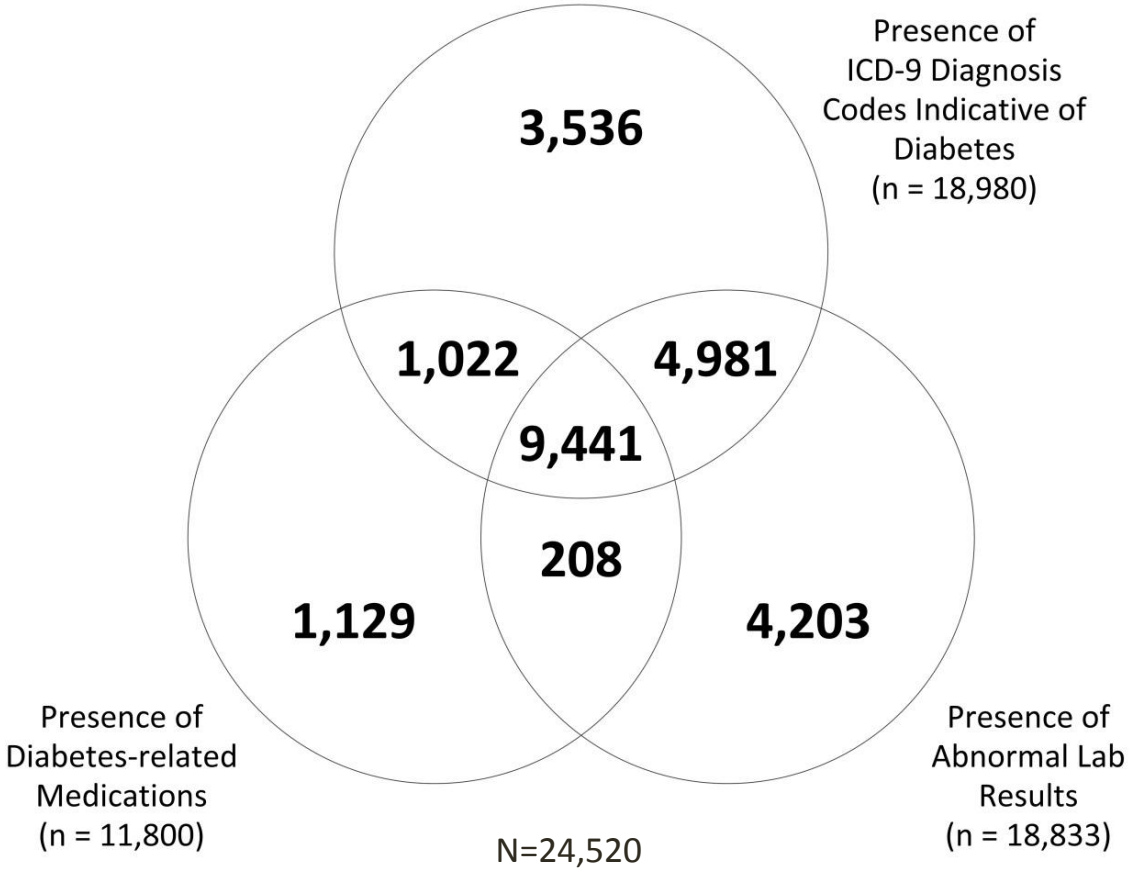
Characteristic	No. (%) of Patients[a]	
	Gentamicin-Collagen Sponge (n = 753)	Control (n = 749)
Patient demographics		
Age, median (IQR), y	64.2 (58.0-71.5)	64.9 (57.2-72.1)
White race	688 (91.4)	683 (91.2)
Weight, median (IQR), kg	98.0 (86.1-113.0)	98.8 (85.0-111.1)
Body mass index, median (IQR)	33.1 (30.2-37.2)	32.8 (30.0-36.2)
Body mass index >30	574 (76.2)	563 (75.2)
Male sex	530 (70.4)	530 (70.8)
Medical history		
History of hypertension	659 (87.5)	659 (88.0)
History of diabetes	493 (65.5)	513 (68.5)
Current or history of smoking	458 (60.8)	450 (60.1)
Current smoking	136 (29.7)	123 (27.3)
History of chronic obstructive pulmonary disease	117 (15.5)	107 (14.3)
History of peripheral vascular disease	105 (13.9)	89 (11.9)
Previous median sternotomy	52 (6.9)	42 (5.6)
History of TIA or stroke	77 (10.2)	81 (10.8)
History of myocardial infarction	233 (31.0)	245 (32.7)
History of congestive heart failure	89 (11.8)	90 (12.0)
History of hyperlipidemia	619 (82.2)	607 (81.0)
Steroid use ≤1 mo prior to surgery	28 (3.7)	33 (4.4)
Receiving dialysis preoperatively	4 (0.5)	2 (0.3)
Preoperative diagnostic values		
Left ventricular ejection fraction, median (IQR), %	55 (45-60)	55 (45-60)
Serum glucose, median (IQR), mg/dL	125 (101-160)	124 (103-167)
Serum hemoglobin A_{1c}, median (IQR), %	6.5 (5.9-7.6)	6.6 (5.9-7.7)
Hematocrit, median (IQR), %	39 (36-42)	39 (36-42)
Serum creatinine, median (IQR), mg/dL	1.0 (0.9-1.3)	1.0 (0.9-1.2)
Preoperative core temperature, median (IQR), °C	97.6 (97.0-98.2)	97.7 (97.0-98.2)
Preoperative hospital stay, median (IQR), d	1.0 (0-3.0)	1.0 (0-3.0)
Parsonnet risk score, median (IQR)[b]	9.0 (6.0-14.5)	9.0 (6.0-16.0)

Abbreviations: IQR, interquartile range; TIA, transient ischemic attack.
SI conversion factors: To convert creatinine to µmol/L, multiply by 88.4; glucose to mmol/L, multiply by 0.0555.
[a] Unless otherwise indicated.
[b] Theoretical range is 0 to 148; 50% in Parsonnet et al[11] had a score between 0 and 9.

Question 3:
Do we want to standardize reporting conventions?
Categories?

e.g.., continuous variables:
- mean and SD
- median, 25th, 75th %-ile
- categories (e.g., deciles of age)

Different Definitions Yield Different Cohorts



Research and applications

A comparison of phenotype definitions for diabetes mellitus

Challenges

- Multiple approaches to defining important characteristics and "standard" presentation:
  - Top-down (expert opinion)
  - Bottom-up (data-driven)

- Condition-specific and study specific components
- Timing of observations/measurements vs. inception of trial
- Observations and fragmentation of care

- Need clearly defined objective
- Need to engage potential users and stakeholders to ensure uptake / future endorsement or adoption

- Connection to pragmatic trials implies EHR as source data

(→ more stakeholders)

Opportunities for the Collaboratory

- Researchers can define best practices in data collection and use
- High-visibility effort – ideal to build (and endorse) a case for standards
- Our members can be a conduit to health care organizations
- We are uniquely focused on getting data from EHRs (in contrast to de novo data collection)

Question: What characteristics should be on Table 1?

- Demographics
  - age, sex, race, ethnicity (federal standards)

- Relevant clinical or behavior/exposure data
  - pan-disease
  - disease-specific

Pan-Disease Features (top half)

- Age
- Gender
- Race
- Ethnicity
- SES
- Height
- Weight
- Blood Pressure
- Insurance Status (to infer access to care)

Condition-Specific (Bottom half)

- MANDATORY:
- Study specific relevant co-morbidities
- Study specific medications
- Study specific labs
- Study specific non-medication interventions
- Cohort ID variables (baseline)

- CAD
- HTN
- Diabetes
- Hyperlipidemia
- CKD
- Anemia
- CHF
- COPD
- Asthma
- PVD
- PUD
- CVA
- Tumor, Leukemia, Lymphoma
- AIDS
- Atrial Arrythmeia
- Dementia
- Connective Tissue Disease
- Cohort identification variables (baseline)

- Co-morbidities - options:
  - Charlson Index
  - Top 10 comorbidities by frequency