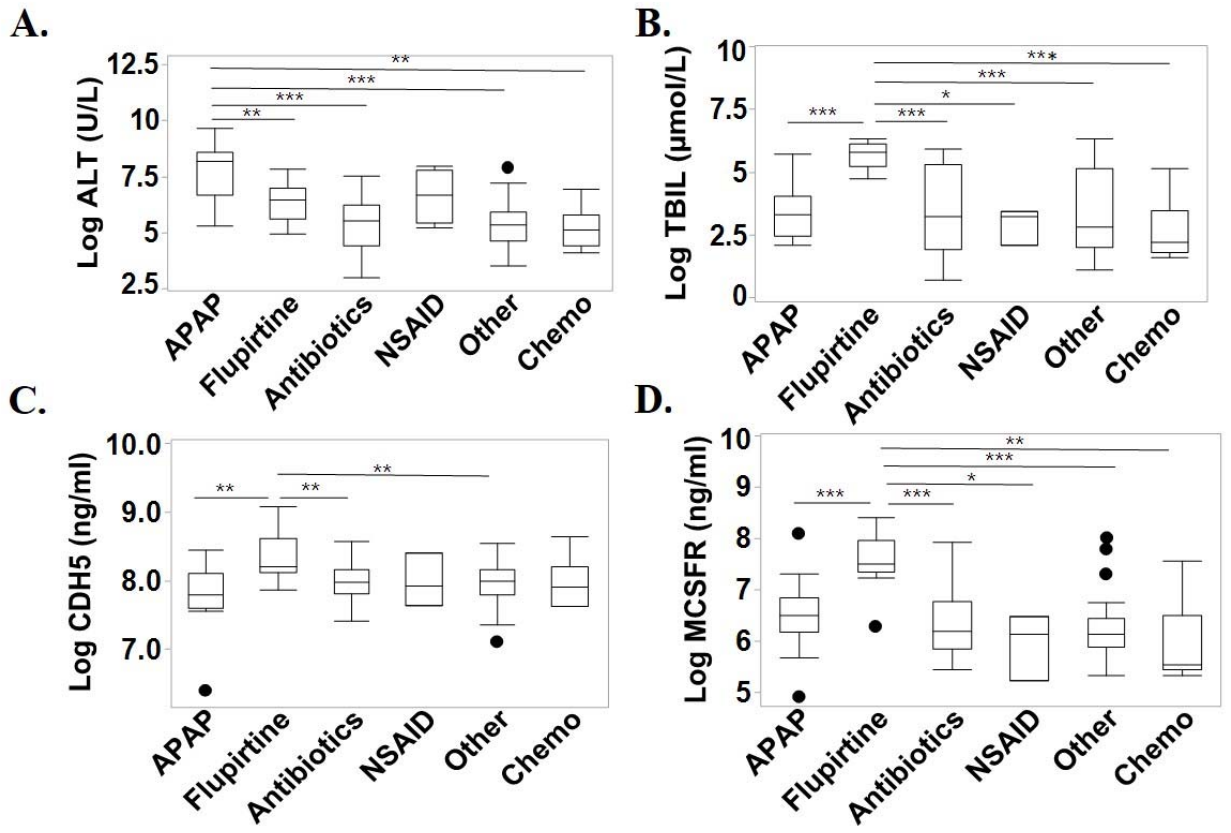
Supplemental Figure 1



Supplemental Figure 1: Biomarker Differences by Drug Class in Safer and Faster Evidence-based Translation (SAFE-T) DILI Patients. Differences in mean alanine aminotransferase (ALT; **A**), total bilirubin (TBIL; **B**), cadherin 5 (CDH5; **C**), and macrophage colony stimulating factor receptor (MCSFR; **D**) between SAFE-T drug-induced liver injury (DILI) patients based on drug classes. Drug classes are acetaminophen (APAP; n=19), flupirtine (n=14), antibiotics (n=35), chemotherapeutics (n=7), non-steroidal anti-inflammatory drugs (NSAIDs; n=4), and others (n=45). The box in each box plot extends from the 25th percentile to the 75th percentile of data values; whiskers extend to minimum and maximum data with data outliers represented by circles. TBIL and CDH5 measurements were collected in serum while MCSFR measurements were collected in plasma. Significance is *p<0.05, **p<0.01, and ***p<0.001.

SUPPLEMENTAL TABLES

Supplemental Table 1: PSTC Demographics

PSTC Healthy Volunteers (N=81)	
Age, y, median (IQR)	39 (29.5-50.5)
Sex, n (%)	
Male	40 (49.4)
Female	41 (50.6)
Race, n (%)	
White	68 (84)
Black	13 (16)
BMI (kg/m^2), median (IQR)	27.8 (23.7-31.35)
Liver biochemistries, median (IQR)	
ALT (U/L)	20 (15.5-28)
AST (U/L)	22 (19-25)
ALP (U/L)	65 (54.5-76.5)
TBIL (µmol/L)	8.55 (6.84-11.97)

Abbreviations: PSTC, Predictive Safety Testing Consortium; IQR, interquatile range; BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; TBIL, total bilirubin

Supplemental Table 2: SAFE-T Demographics

	SAFE-T Study				
	Healthy Volunteers	**No DILI**		**DILI**	
	Tel Aviv	**Protocol 4**	**Protocol 5**	**Swiss DILI**	**Protocol 3A**
	N=192	**N=55**	**N=92**	**N=28**	**N=98**
Age, y, median (IQR)	52 (42-62)	29 (24-39)	52.5 (43.3-61)	56 (42-66.8)	53 (38-66.3)
Sex, n (%)					
Male	103 (53.7)	32 (58.2)	31 (33.7)	15 (53.6)	41 (41.8)
Female	88 (45.8)	23 (41.8)	61 (66.3)	13 (46.4)	9 (47.4)
Missing	1 (0.5)				
Race, n (%)					
White		33 (60)	68 (73.9)	25 (89.3)	90 (91.8)
Black		20 (36.4)	11 (12)	2 (7.1)	1 (1.02)
Asian		1 (1.8)	1 (1.1)	0 (0)	5 (5.1)
Other		1 (1.8)	10 (10.8)	1 (3.6)	1 (5.3)
Missing			2 (2.2)		
BMI (kg/m^2), median (IQR)	25.9 (23.1-29.2)	21.6 (19.3-24.8)	25.2 (22-29.4)	24.6 (22-27.4)	25.7 (23.4- 29)
Liver biochemistries, median (IQR)					
ALT (U/L)	22 (18-29)	21.5 (18-35)	25 (18-32)	278 (144-1877)	322 (137.8-884)
AST (U/L)	23 (20-26)	26 (20.8-33)	26.5 (22-30)	152 (64-728)	138.6 (66.5-349)
ALP (U/L)	66 (54-81)	71 (57-9.8)	62 (46-73.5)	84.5 (65-246.8)	181 (101- 254)
TBIL (µmol/L)	10.26 (8.6-13.7)	6.8 (5.1-10.3)	7 (5-9.8)	8.5 (5.3-32)	42 (11.5-247)
INR				1.1 (1-1.3)	1.3 (1-1.6)
Hy's Law, n (%)					
No				20 (71.4)	53 (54.1)
Yes				4 (14.3)	35 (35.7)
Missing				4 (14.3)	10 (10.2)
Pattern of Injury, n (%)					
Cholestatic				6 (21.4)	5 (5.1)
Mixed				1 (3.6)	24 (24.5)
Hepatocellular				21 (75)	69 (70.4)
R Value, median (IQR)				7.6 (1.3-75)	5.7 (2.3- 27.9)

Abbreviations: SAFE-T, Safer and Faster Evidence-based Translation; DILI, drug-induced liver injury; IQR, interquartile range; BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; TBIL, total bilirubin; INR, international normalized ratio; Hy's Law (ALT>3X Upper Limit of Normal, ULN, TBIL>2X ULN, ALP<2X ULN)

Supplemental Table 3: DILIN Demographics

	DILI Outcome				
	Recovered	**Unresolved**	**Death/Transplant**	**Unknown**	***p***
	N=89	**N=19**	**N=15**	**N=20**	
Age, y, mean ± S.D	46.2 ± 16.9	43.1 ± 16.1	52.6 ± 20.8	39.9 ± 15.9	NS
Sex, n (%)					NS
Male	32 (36)	10 (52.6)	9 (60)	11 (55)	
Female	57 (64)	9 (47.4)	6 (40)	9 (45)	
Race, n (%)					NS
White	65 (73)	12 (63.2)	10 (66.7)	16 (80)	
Black	13 (14.6)	6 (31.5)	2 (13.3)	3 (15)	
Asian	5 (5.6)	0 (0)	2 (13.3)	0 (0)	
Other	6 (6.8)	1 (5.3)	1 (6.7)	1 (5)	
Ethnicity, n (%)					NS
Hispanic	13 (14.6)	2 (10.5)	0 (0)	3 (15)	
Non-Hispanic	76 (85.4)	17 (89.5)	15 (100)	17 (85)	
BMI (kg/m^2), mean ± S.D.	28.5 ± 7.3	28.1 ± 9.6	27.4 ± 6.6	26.5 ± 5.0	NS
Liver biochemistries, median (IQR)					
ALT (U/L)	527 (228.8-1258.5)	357 (128-1106)	907 (152-1536)	247 (106-458.3)	NS
AST (U/L)	306 (126.3-755.3)	290 (71-664)	865 (220-987)	130 (63.25-612.3)	0.01
ALP (U/L)	165 (127.3-323.5)	216 (173- 327)	146 (120-297)	229.5 (152.3-356.8)	NS
TBIL (µmol/L)	93.2 (26.5-221.9)	165.9 (73.5-311.2)	311.2 (261.6-434.3)	177.8 (47.9-262.5)	<0.0001
INR	1.1 (1-1.3)	1.1 (1-1.4)	3 (1.7-4.4)	1.0 (0.9-1.2)	<0.0001
Hy's Law, n (%)					0.007
No	54 (61.4)	12 (63.2)	3 (20)	15 (75)	
Yes	34 (38.6)	7 (36.8)	12 (80)	5 (25)	
Pattern of Injury, n (%)					NS
Cholestatic	17 (19.3)	8 (42.1)	3 (20)	10 (50)	
Mixed	16(18.2)	2 (10.5)	3 (20)	4 (20)	
Hepatocellular	55 (62.5)	9 (47.4)	9 (60)	6 (30)	
R Value, median (IQR)	8.2 (2.3-19.9)	3.8 (1- 14.7)	13.7 (3.2-36.6)	2.2 (0.9-9.3)	NS
MELD Score, median (IQR)	16.1 (103-21.7)	16.7 (12.2-19.1)	33.2 (28.9-40)	17.4 (12.9-20.2)	<0.0001

Abbreviations: DILI, drug-induced liver injury; BMI, body mass index; IQR, interquartile range; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; TBIL, total bilirubin; INR, international normalized ratio; Hy's Law (ALT>3X Upper Limit of Normal, ULN, TBIL>2X ULN, ALP<2X ULN); MELD, Model of End-stage Liver Disease

Supplemental Table 4: Primary Causative Drugs in DILI Patients

Primary Causative Drug	Patients (n)		Drug Class
	SAFE-T	DILIN	
Acetaminophen	19		APAP
Acetaminophen and Others	3		APAP + Others
Acetazolamide	1		Others
Allopurinol		1	Others
Althiazide	1		Others
Amino Acids Nos		1	Others
Amiodarone		1	Others
Amoxicillin	1	1	Antibiotics
Amoxicillin W/Clavulanic Acid	9	11	Antibiotics
Amphetamines	1		Others
Anabolic Agents For Systemic Use	3	4	Others
Anakinra		1	Others
Antiinflammatory And Antirheumatic Products,		1	Others
Antithymocyte Immunoglobulin		1	Others
Asparaginase		1	Chemotherapy
Atorvastatin	5		Others
Azathioprine	1	1	Others
Azithromycin	2	3	Antibiotics
Baclofen	1		Others
Beta-Interferon	2		Others
Bortezomib	1		Others
Bupropion		1	Others
Camellia Sinensis		1	Others
Carbamazepine	1	2	Others
Carbohydrates/Proteins/Minerals/Vitamins, Com		1	Others
Cefalexin		1	Antibiotics
Cefazolin		2	Antibiotics
Cefotaxime		1	Antibiotics
Ceftriaxone	2	1	Antibiotics
Celecoxib	1		NSAID
Centrally Acting Sympathomimetics		1	Others
Ciprofloxacin		4	Antibiotics
Clarithromycin		1	Antibiotics
Cyclophosphamide	1		Chemotherapy
Cyclosporine A	1		Others
Dantrolene		1	Others
Dapsone		1	Antibiotics
Daptomycin		1	Antibiotics
Darunavir		2	Others
Diclofenac		2	NSAID
Disulfiram	1	1	Others
Doxepin	1		Others
Doxycycline	1	1	Antibiotics
Erythromycin W/Sulfisoxazole		1	Antibiotics
Escitalopram		2	Others

Primary Causative Drug	Patients (n)	Drug Class	Primary Causative Drug
Etoricoxib	1		NSAID
Exemestane	1	1	Others
Fenofibrate	1	2	Others
Fingolimod	1		Others
Flavocoxid		1	Others
Flucloxacillin	6		Antibiotics
Flupirtine	14		Flupirtine
Fluvastatin	1		Others
Gabapentin	1		Others
Herbals		7	Others
Hydralazine		2	Others
Hydroxycut - Ephedra Free		2	Others
Ibuprofen	2		NSAID
Imetelstat		1	Chemotherapy
Infliximab		2	Others
Ipilimumab		1	Chemotherapy
Isoniazid		12	Anti-TB
Isoniazid/pyrazinamide/rifampin	1		Anti-TB
Isoniazid/pyrazinamide/rifampin/ ethambutol	1		Anti-TB
Leflunomide	2		Others
Letrozole	1		Others
Levofloxacin	1	3	Antibiotics
Lisinopril		1	Others
Metamizole	3		Others
Mercaptopurine		2	Chemotherapy
Meropenem	1		Antibiotics
Methyldopa	1	2	Others
Micafungin		1	Others
Minocycline		5	Antibiotics
Montelukast		1	Others
Moxifloxacin		1	Antibiotics
Mushrooms	1		Others
Nefazodone		1	Others
Nicotinic Acid		3	Others
Nitrofurantoin	1	2	Antibiotics
Octreotide		1	Others
Olanzapine	1		Others
Oxaliplatin		2	Chemotherapy
Oxymethalone	1		Others
Pantaprazole	1		Others
Pentamidine	1		Others
Pegaspargase		2	Chemotherapy
Phenprocoumon	1		Others
Phenylpropanolamine		1	Others
Phenytoin		3	Others
Piperacillin Sodium W/Tazobactam	2	1	Antibiotics
Pravastatin		1	Others

Primary Causative Drug	Patients (n)	Drug Class	Primary Causative Drug
Prednisolone	1		Others
Pregabalin	1	1	Others
Propylthiouracil		2	Others
Quetiapine		2	Others
Rifampin	1		Anti-TB
Several Antibiotics	5		Antibiotics
Several Chemoterapeutics	5		Chemotherapy
Simvastatin	1	1	Others
Sulfamethoxazole W/Trimethoprim	1	11	Antibiotics
Sulfasalazine		1	Others
Tacrolimus	1		Others
Temozolomide	1		Chemotherapy
Terbinafine	1		Others
Thiamazole	1		Others
Valaciclovir		1	Others
Valproic Acid		1	Others
Other		3	Others
Query Outstanding		1	

Abbreviations: SAFE-T, Safer and Faster Evidence-based Translation; DILIN. Drug-Induced Liver Injury Network; APAP, acetaminopen

Supplemental Table 5: Biomarker Validation Data

Analyte	Type of Assay	Sample Matrix analyzed	unit	LOD	LLoQ	ULoQ	intra-assay precision (% CV)	inter-assay precision (% CV)	dilutional linearity of high conc sample	Spike-in recovery (%)	short term stability (24h at RT and 4°C)	F/T stability, 3 cycles
ccK18	ELISA	Serum	U/L	16.2	62.5	1000	2.2	5.7 - 7.9	up to 1:16	112 - 118	yes	yes
K18	ELISA	Serum	U/L	20	100	5000	3.7	6.1 - 9.4	up to 1:32	83 - 107	yes	yes
GLDH	Activity Assay	Serum	U/L	0.3	1	80	0.4 - 7.7	1.5 - 6.4	1:4 - 1:256	ND	yes, > 6h	yes
GSTα	Immunoassay	Serum	ng/mL	1.79	1.82	373	1 - 14	11-Sep	1:5 - 1:10	77 - 94	yes	yes
AFP	Immunoassay	Serum	ng/mL	0.367	0.367	584	16-Feb	13-Jul	1:5 - 1:40	99 - 106	yes	yes
ARG1	Immunoassay	Serum	ng/mL	1.6	7.4	800	6.4 - 11.9	4.3 - 15.7	1:4 - 1:256	84 - 88	yes	yes
OPN	Immunoassay	Serum	ng/mL	1.25	1.25	1149	5-Jan	6 - 11	1:5 - 1:10	81 - 85	yes	yes
SDH	Activity Assay	Serum	U/L	0.3	0.5	50	0.6 - 10.6	1.7 - 13.4	up to 1:32	ND	yes, > 6h	yes
miR-122	RT-qPCR	Serum	copies/µL	ND	384	5089837	1.3 - 12.1	0.5 - 25.4	ND	ND	2h RT, 5h 4°C	yes
FABP1	Immunoassay	Serum	pg/mL	3.1	15.6	16000	5.6	6.7 - 18.1	1:2 - 1:2048	110 - 115	yes	yes
CDH5	ELISA	Serum	ng/mL	0.36	3.13	100	6	4.7 - 7.2	1:40 - 1:640	50 - 83	yes	yes
MCSFR	Immunoassay	EDTA-Plasma	pg/mL	170	600	10000	1.1 - 13.9	8.0 - 28.0	up to 1:3,200	71 -79	yes	yes
PON1	Immunoassay	EDTA Plasma	ng/mL	0.06	0.35	600	5.9	8.3 - 12.3	1:20 - 1:160	64 - 82	4h RT, 24h 4°C	yes
Prothrombin	Immunoassay	EDTA Plasma	µg/mL	0.8	1.92	200	4.7	1.7 - 4.5	1:40 - 1:320	79 - 108	yes	yes
LECT2	Immunoassay	EDTA Plasma	ng/mL	2	5.56	300	7.8	11.7 - 12.6	1:40 - 1:1.280	94 - 118	yes	yes

Abbreviations: limit of detection (LOD) , lower limit of quantification (LLoQ),upper limit of quantification (ULoQ), coefficiant of variability (CV), concentration (conc), hours (h), room termperature (RT), freeze/thaw (F/T), total cytokeratin 18 (K18), caspase cleaved cytokeratin 18 (ccK18), glutamate dehydrogenase (GLDH), not determined (ND), glutathione-S-transferase α (GSTα), alpha fetoprotein (AFP), arginase 1 (ARG1), osteopontin (OPN), sorbitol dehydrogenase (SDH), microRNA-122 (miR-122), reverse transcription quantitative real-time PCR (RT-qPCR), liver fatty acid binding protein (L-FABP), cadherin 5 (CDH5), macrophage colony stimulating factor receptor (M-CSF-R), paroxonase 1 (PON1, normalized to prothrombin protein), leukocyte cell-derived chemotaxin 2 (LECT2)

Supplemental Table 6: Biomarker Alterations in Augmentin-related DILI			
Biomarker	**Mean Biomarker Value (Ln)**		***p***
	SAFE-T	**DILIN**	
ALT (U/L)	4.67	5.46	0.048
ARG1 (ng/ml)	2.83	3.52	0.033
FABP1 (ng/ml)	2.82	3.82	0.04
ccK18 (U/L)	5.75	6.29	0.028

Abbreviations: DILI, drug-induced liver injury; SAFE-T, safer and faster evidence-based translation; DILIN, DILI network; ARG1, arginase 1; FABP1, fatty acid binding protein 1; ccK18, caspase cleaved keratin 18

Supplemental Table 7: Biomarker Geometric Means of Healthy Volunteers and DILIN Patients

Biomarker	Geometric Mean			Fold Change	
	HV	No Death/Trans	Death/Trans	HV vs. Death/Trans	No Death/Trans vs. Death/Trans
OPN (ng/ml)	5.75	14.17	41.01	7.13	2.89
K18 (U/L)	68.44	1358.73	10481.29	153.15	7.71
MCSFR (ng/ml)	315.4	883.93	2240.95	7.11	2.54
ccK18 (U/L)	121.83	978.14	3636.49	29.85	3.72
FABP1 (ng/ml)	8.54	50.14	133.7	15.66	2.67
AFP (ng/ml)	0.9	4.47	10.32	11.47	2.31

Abbreviations: DILIN, drug-induced liver injury network; HV, healthy volunteer; trans, transplant; OPN, osteopontin; K18, total keratin 18; MCSFR, macrophage colony stimulating factor receptor; ccK18, caspase cleaved keratin 18; FABP1, fatty acid binding protein 1; AFP, alpha fetoprotein

Supplemental Table 8: Comparison of Prediction Models for Death/Liver Transplant

Model	Sensitivity	Specificity	PPV	NPV	*p*
Hy's Law	0.8	0.634	0.207	0.964	0.0054
MELD score ≥ 20	0.933	0.738	0.298	0.989	<0.0001
MELD score ≥ 30	0.6	0.992	0.9	0.954	<0.0001
Modified Hy's Law*	0.733	0.611	0.183	0.951	0.0303
ALF Algorithm*	0.533	0.817	0.258	0.936	0.0075
MELD + K18/MCSFR	0.933	0.889	0.5	0.991	<0.0001

Abbreviations: PPV, positive predictive value; NPV, negative predictive value; Hy's Law (ALT>3X Upper Limit of Normal, ULN, TBIL>2X ULN, ALP<2X ULN); MELD, Model of End-stage Liver Disease; ALF, acute liver failure; K18, total keratin 18; MCSFR, macrophage colony stimulating factor receptor.

*Robles-Diaz M, Lucena MI, Kaplowitz N, Stephens C, Medina-Caliz I, Gonzalez-Jimenez A, *et al.* Use of Hy's law and a new composite algorithm to predict acute liver failure in patients with drug-induced liver injury. Gastroenterology 2014;147:109-118 e5.

Supplemental Table 9: Prognostic Biomarkers for Unresolved DILI at 6 Months Post-Onset

Category	Biomarker	AUC	95% CI
Traditional	ALP	0.67	0.562-0.777
Traditional	TBIL	0.629	0.497-0.761
Traditional	ALT	0.544	0.39-0.7
Traditional	INR	0.528	0.376-0.679
Traditional	AST	0.516	0.369-0.664
Candidate	GST-α	0.633	0.485-0.78
Candidate	ARG1	0.614	0.48-0.747
Candidate	ccK18	0.58	0.442-0.719
Candidate	OPN	0.562	0.436-0.688
Candidate	FABP1	0.562	0.418-0.706
Candidate	CDH5	0.539	0.406-0.673
Candidate	AFP	0.538	0.397-0.679
Candidate	K18	0.519	0.374-0.664
Candidate	MCSFR	0.516	0.385-0.647
Candidate	AI	0.53	0.359-0.702

*All values with the exception of AI are log normalized

Abbreviations: AUC, area under the curve; CI, confidence interval; INR, international normalized ratio; TBIL, total bilirubin; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALP, alkaline phosphatase; OPN, osteopontin; K18, cytokeratin 18; MCSFR, macrophage colony stimulating factor receptor; ccK18, caspase cleaved K18; FABP1, fatty acid binding protein 1; AFP, alpha fetoprotein; ARG1, arginase 1; CDH5, cadherin 5; GST-α, glutathione S transferase alpha; AI, apoptotic index

Supplemental Table 10: Biomarker Correlation with Days Between Symptom Onset and Biospecimen Collection in DILIN Patients

Biomarker	Pearson's r
AFP	0.077
ARG1	-0.014
CDH5	0.222
K18	0.142
ccK18	0.212
FABP1	-0.002
GST-α	-0.048
MCSFR	0.163
OPN	0.158

Abbreviations: DILIN, drug-induced liver injury network AFP, alpha fetoprotein; ARG1, arginase 1; CDH5, cadherin 5; K18, cytokeratin 18; ccK18, caspase cleaved K18 FABP1, fatty acid binding protein 1; GST-α, glutathione S transferase alpha; MCSFR, macrophage colony stimulating factor receptor; OPN, osteopontin;

SUPPLEMENTAL METHODS

PSTC Healthy Volunteers

All subjects in this cohort (n=81; **Supplemental Table 1**) were recruited at the Jasper Clinic, Inc., Kalamazoo, MI, USA. Three fasting blood samples (n=243 total samples) were collected from 81 subjects over 21 days. Inclusion criteria included age between 18 and 70 years, no underlying medical conditions or use of chronic medications, and a body mass index (BMI) < 35 (kg/m^2) (two exceptions with BMIs of 35.3 and 37.6 kg/m^2). Exclusion criteria included a positive test for human immunodeficiency virus, and/or active hepatitis B or hepatitis C viral infections, a medical intervention performed within three months of study enrollment, a positive pregnancy test, or unwillingness to refrain from illicit drug/alcohol/tobacco use or strenuous exercise during the study.

SAFE-T Healthy Volunteers

Biomarker measurements from subjects in this cohort (n=192; **Supplemental Table 2**) were taken from a single fasting blood sample collected between 7 and 9 a.m. at the Tel Aviv Sourasky Medical Center, Tel Aviv Israel. All subjects were asymptomatic and in good health. They completed a detailed epidemiological questionnaire and underwent a thorough analysis of life style by a trained nutritionist. Subjects were interviewed regarding their personal and family history and underwent a comprehensive physical examination. Female subjects underwent a breast and pelvic exam by a senior surgeon and mammography was performed at age > 40 years. Heavy smokers (>20 packs/year) were offered a computed tomography scan. Men > 40 years were tested for total and free prostate-specific antigen. Further diagnostic tests were performed as needed based on the

initial screening results. Exclusion criteria included heavy alcohol intake, a history of renal or liver diseases, and a personal or family history of cancer.

SAFE-T DILI Patients

The clinical studies analyzed in this manuscript can be divided into (i) protocols that recruited patients diagnosed with DILI (**Supplemental Table 2**, "DILI") and (ii) protocols that recruited patients who safely took a known DILI-eliciting compound and who were prospectively monitored for several months without evidence of liver injury (**Supplemental Table 2**, "No DILI"). Fasting blood samples were collected. The SAFE-T criteria for adjudicating suspected DILI cases have been described elsewhere [1]. With few exceptions, DILI patients fulfilled the consensus criteria for DILI as previously published [2, 3].

SAFE-T DILI Patients:

Swiss DILI study: This study was an 8-week single-center follow-up study investigating the prognostic value of new biomarkers in patients with DILI and included 28 patients adjudicated as having DILI. None of the patients included from this protocol died/required a liver transplant during the observation period. It is unknown whether these patients developed chronic DILI.

Protocol 3A: This study was a 12-week multi-center follow-up study investigating the prognostic value of new biomarkers in patients with DILI and included 98 patients adjudicated as DILI. None of the patients included from this protocol died/required a liver transplant during the observation period. It is unknown whether these patients developed chronic DILI.

SAFE-T Drug-exposed No DILI Patients:

Protocol 4: This study was a 9-month single-center follow-up study in tuberculosis patients (n=55) starting anti-tuberculosis drug therapy. None of the patients enrolled in this protocol developed

DILI [ALT >5X upper limit of normal (ULN)] during the observation period. Biomarker measurements were made in samples collected at a time point after the patients had begun taking compound (time ranged from 1-6 months on compound).

Protocol 5: This study was a 3-year single-center follow-up study in rheumatoid arthritis patients and 92 patients were included in this analysis. None of the patients enrolled in this protocol developed DILI (ALT >5X ULN) during the observation period. When possible biomarker measurements were made in samples collected at time points after patients had begun taking compound (time ranged from 6-30 months on compound); however, only a baseline sample was available for some of these individuals (n=26).

DILIN Patients

DILIN prospectively collects clinical, laboratory, imaging, and histopathological information as well as biospecimens from patients within 6 months of suspected DILI onset at multiple centers across the United States (**Supplemental Table 3**). The criteria utilized for DILI assessment in this network has been described in detail elsewhere [4]. The current study assessed biomarkers in 143 samples and included only patients with probable, highly likely, or definite DILI and a blood sample collected within two weeks of DILI onset. Within this cohort, 15 patients died/required a liver transplant within 6 months of onset because of their DILI. Following a readjudication process, DILI was deemed to be the primary factor in all of these patients [5]. Additionally, 19 patients had unresolved DILI (persistently elevated ALT, AST, ALP, or TBIL in the absence of a competing etiology) at 6 months following onset. Of the remaining patients, 89 had recovered by their 6

months follow up visit and 20 did not return for a follow up visit, therefore it is unknown whether their liver injury had completely resolved.

Biomarker Quantification

Predictive Safety Testing Consortium (PSTC) and Safer and Faster Evidence-based Translation (SAFE-T) biomarker measurements were made in either serum or EDTA-plasma (plasma) depending on which matrix was determined to be better suited for the assay. Leucocyte cell derived chemotaxin 2 (LECT2), macrophage colony stimulating factor receptor (MCSFR), and paraoxonase 1 (PON1; normalized to prothrombin protein) were quantified in plasma. All other biomarkers were quantified in serum. All Drug-Induced Liver Injury Network (DILIN) biomarker measurements were made in serum samples. Of the subset of biomarkers measured in all datasets (due to limitations on sample volume, only 9/14 biomarkers were examined in DILIN patients), the matrix for MCSFR differed between cohorts because of sample availability. For all analytes, no international reference standard was available and the measured concentrations were calculated based on individual standard proteins used for the assay calibration.

Traditional biomarkers alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), total bilirubin (TBIL), and international normalized ratio (INR) were measured at the local institutional clinical laboratories and were not obtained from stored samples. Samples utilized for candidate biomarker analyses were taken from archived samples stored at ≤ -70°C. Measurements were made at Natural and Medical Sciences Institute (NMI; Reutlingen, Germany) or at contract research laboratories. Briefly, ELISA assays were used to measure total keratin 18 (K18), caspase cleaved K18 (ccK18; VLVbio, Stockholm, Sweden) and cadherin 5 (CDH5; R&D Systems, Minneapolis, MN). Sandwich immunoassays were used to measure glutathione S transferase alpha (GSTα), alpha fetoprotein (AFP), osteopontin (OPN; optimized

Luminex assays from Myriad RBM, Austin, Texas) and arginase 1 (ARG1), MCSFR, PON1, prothrombin protein, fatty acid binding protein 1 (FABP1), and LECT2 (assays developed by NMI, Germany). Colorimetric applications for glutamate dehydrogenase (GLDH) activity (Roche Diagnostics, Grenzach-Wyhlen, Germany) and sorbitol dehydrogenase (SDH) activity (Sekisui Diagnostics, Lexington, MA, USA) were run on a Roche P. Modular Analyzer. PON1 was normalized to prothrombin protein because evidence suggests that this normalization method enables distinction from nonacloholic steatohepatitis and nonalcoholic fatty liver disease [6]. Absolute quantification of microRNA-122 (miR-122) was analyzed by reverse transcription quantitative real time PCR utilizing standard reagents and real time hydrolysis probes (Life Technologies, Grand Island, New York). Differences in RNA extraction efficiency from individual serum samples were compensated for by adding a synthetic non-human miR (mmu-miR-293) to all samples prior to extraction. All PCR analyses were performed on 192.24 Dynamic Array IFC (Fluidigm). Cq values were calculated by averaging the technical triplicate Cq values, normalized by the average Cq value of the spiked mmu-miR-293 and total miR-122 copy numbers/μL were calculated.

When a biomarker value fell below the lower limit of quantification (LLoQ), that value was used as LLoQ/2.

All commercial assay kits were run according to manufacturer's recommended protocols. All non-clinical assays used for analysis of sample sets which were performed at NMI or contract research organizations were validated following a fit-for-purpose approach considering usual guidelines.

Validation of each assay was approved by a dedicated team within the SAFE-T consortium before the assays were released for sample screenings.

When permissable, an apoptotic index of injury (AI) was calculated from patient data utilizing the ratio of ccK18 to K18. Evidence has demonstrated that this ratio is only meaningful when ccK18 and K18 are above background threshhold levels [7]. In the current study, the following rules were set to establish when calculation of an AI was appropriate: a) K18 ≥ 500 U/L b) ccK18 ≥ 200 U/L c) K18 > ccK18. Using these rules, an AI was calculated for 98 DILIN patients and 64 SAFE-T DILI patients. Significance was determined by logistic regression and was considered $p<0.05$.

Biomarker Differences by Drug Class

To determine if one or more DILI compounds/classes produces signature biomarker changes that are unique compared to APAP-related DILI, SAFE-T DILI patient data were divided into broad drug classes. Data was divided as follows: APAP (n=19), flupirtine (n=14), antibiotics (n=35), chemotherapeutics (n=7), non-steroidal anti-inflammatory drugs (NSAIDs; n=4), and others (n=45). When a primary causative drug was uncertain, data were excluded (n=2). Biomarker differences in drug classes were determined in SAFE-T DILI data and DILIN patient data utilizing a one way ANOVA and Wilcoxon multiple comparison correction.

Additionally, cohort differences in patients with DILI related to amoxicillin with clavulanic acid (Augmentin) was examined between DILIN (n=11) and SAFE-T (n=9). Differences were determined using a Wilcoxon test.

Prognostic Model Generation

The performance of current DILI outcome prediction models including Hy's Law, Model for End Stage Liver Disease (MELD) ≥ 20, MELD ≥ 30, along with a modified version of Hy's Law and

a novel model proposed to predict acute liver failure in DILI patients [8] were explored in the current DILIN patient cohort. Patients were assigned a binary label based on whether or not they met model criteria. Hy's Law criteria was met if patients had ALT ≥ 3X upper limit of normal (ULN), TBIL ≥ 2X ULN, and ALP < 2X ULN. A MELD score for each patient was calculated as previously described [9]. Concurrent sodium levels were not utilized in this calculation. Performance characteristics (sensitivity, specificity, positive predictive value, PPV, and negative predictive value, NPV) were determined and a contingency table and Fisher's exact test were used to establish significance.

We were interested in determining if candidate DILI biomarkers added value to predictions of death/transplant made with traditional biomarkers. The statistical literature and related data suggest that at least 10 cases are needed for every covariate in a logistic regression prediction model to avoid over-fitting; otherwise, the parameter estimators will be unstable, the covariates in the model may represent noise instead of the true effects of underlying risk factors, and precision of parameter estimators will be poor [10]. Because only n=15 patients in this cohort required a liver transplant or died as a result of DILI, construction of a predictive model using only the biomarker data from this study was not attempted. Instead, we sought to determine if incorporation of any candidate biomarkers could improve the performance of common or previously described predictive models (that use traditional biomarker data). To reduce the number of candidate biomarkers being examined in this analysis, only biomarkers considered predictive of outcome (AUC and lower tail of 95% CI both > 0.5) were carried forward. Predictive biomarkers were then used to construct a correlation matrix and Pearson's r for each biomarker combination was

determined. If any biomarkers were found to be highly correlated, only the biomarker with the greatest AUC generated in ROC curve analysis was carried forward.

Novel biomarkers were also incorporated into a model that utilized MELD score, given that MELD ≥ 20 was the most sensitive prediction model and MELD ≥ 30 was the most specific prediction model. Because most of the ”false” tests when using MELD score were observed when a patient’s MELD score was between 20 and 30, we determined if adding novel biomarker quantifications to this subset of patients would improve the MELD performance. Any patient with a MELD score <20 was considered ”recovered.” And patient with a MELD score ≥30 was considered ”adverse.” Using the biomarkers that passed our earlier filters (K18, OPN, MCSFR, and AFP) we first determined the single biomarker that best impoved the specifity of the MELD score model (one ”adverse” patient had a MELD score <20, therefore the sensitivity of the model could not be further improved using this approach) without negatively affecting the sensitivity. Once this biomarker was identified, we determined if adding a second biomarker could improve the specificity further.

This analysis was performed using data from 141 DILIN patients (2 patient had missing laboratory values). Youden’s J is a single statistic that estimates the probability of an informed decesion and captures the performance of a binary test. Therefore, The value corresponding to the best Younden’s J for each biomarker was used as an unbiased cut-off threshold for calling the outcome of each patient. Receiver operating characteristic (ROC) curve performance characteristics were examined when each biomarker was added alone or when a combination of the candidate biomarkers was incorporated. The combination of biomarkers that gave the best performance (K18 and MCSFR) was reported.

Supplemental References

1. Mikus M, Drobin K, Gry M, Bachmann J, Lindberg J, Yimer G, *et al.* Elevated levels of circulating CDH5 and FABP1 in association with human drug-induced liver injury. Liver Int 2016;
2. Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, *et al.* Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther 2011;89:806-15.
3. Fontana RJ, Watkins PB, Bonkovsky HL, Chalasani N, Davern T, Serrano J, *et al.* Drug-Induced Liver Injury Network (DILIN) prospective study: rationale, design and conduct. Drug Saf 2009;32:55-68.
4. Fontana RJ, Hayashi PH, Gu J, Reddy KR, Barnhart H, Watkins PB, *et al.* Idiosyncratic drug-induced liver injury is associated with substantial morbidity and mortality within 6 months from onset. Gastroenterology 2014;147:96-108 e4.
5. Hayashi PH, Rockey DC, Fontana RJ, Tillmann HL, Kaplowitz N, Barnhart HX, *et al.* Death and liver transplantation within 2 years of onset of drug-induced liver injury. Hepatology 2017;66:1275-1285.
6. Bell LN, Theodorakis JL, Vuppalanchi R, Saxena R, Bemis KG, Wang M, *et al.* Serum proteomics and biomarker discovery across the spectrum of nonalcoholic fatty liver disease. Hepatology 2010;51:111-20.
7. Linder S, Olofsson MH, Herrmann R, and Ulukaya E. Utilization of cytokeratin-based biomarkers for pharmacodynamic studies. Expert Rev Mol Diagn 2010;10:353-9.
8. Robles-Diaz M, Lucena MI, Kaplowitz N, Stephens C, Medina-Caliz I, Gonzalez-Jimenez A, *et al.* Use of Hy's law and a new composite algorithm to predict acute liver failure in patients with drug-induced liver injury. Gastroenterology 2014;147:109-118 e5.

9. Kamath PS and Kim WR. The model for end-stage liver disease (MELD). Hepatology 2007;45:797-805.

10. Peduzzi P, Concato J, Kemper E, Holford TR, and Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373-9.