

## Supplementary Online Content

Ley C, Heath F, Hastie T, Gao Z, Protsiv M, Parsonnet J. Defining usual oral temperature ranges in outpatients using an unsupervised learning algorithm. *JAMA Intern Med*. Published online September 5, 2023. doi:10.1001/jamainternmed.2023.4291

**eTable 1.** *ICD-10* Codes and Medications Excluded by LIMIT, by Frequency

**eTable 2.** Encounters That Did Not Meet Initial Inclusion Criteria Into the Pre-LIMIT Sample

**eTable 3.** Most Common *ICD-10* Codes ( $\geq 0.3\%$ ) in the Analysis Set, by Frequency

**eTable 4.** Linear Regression and Mixed-Effects Modeling of Temperature as a Function of Time of Day, Age, Height, Weight and Month, in Men and Women

**eAppendix.** Missing Data—Weight and Height

**eFigure.** View of the Interactive Visualization Tool for Individualized Normal Temperature Ranges With Examples

**eMethods.** Data and Code (SAS and R)

This supplementary material has been provided by the authors to give readers additional information about their work.

**eTable 1. ICD-10 Codes\* and Medications Excluded by LIMIT, by Frequency**

Diagnosis (ICD-10 code) or medication	Dropped instances		Low temperature outliers		High temperature outliers	
	N	% <sup>d</sup>		% <sup>d</sup>	N	% <sup>d</sup>
Type 2 diabetes mellitus (E11) <sup>a, b</sup>	57,179	15.71	968	26.26	395	2.87
Cough (R05)	31,771	8.73	273	7.41	1,008	7.32
Macrolides	27,101	7.45	213	5.78	1,256	9.12
Beta-adrenergic agents, inhaled, short acting	25,791	7.09	224	6.08	740	5.37
Acute upper respiratory disease (J6)	25,293	6.95	196	5.32	881	6.4
Penicillins	19,219	5.28	177	4.8	801	5.82
Acute sinusitis (J01)	15,759	4.33	158	4.29	414	3.01
Acute pharyngitis / sore throat (J02)	14,299	3.93	114	3.09	782	5.68
Quinolones	13,732	3.77	121	3.28	662	4.81
Acute bronchitis (J20)	12,565	3.45	89	2.41	595	4.32
Disease of urinary tract (N39) <sup>a</sup>	11,568	3.18	125	3.39	208	1.51
Narcotic antitussive-expectorant combination	10,946	3.01	65	1.76	542	3.94
Narcotic antitussive-1 <sup>st</sup> generation antihistamine	9,268	2.55	97	2.63	560	4.07
Antivirals, general	8,964	2.46	70	1.9	533	3.87
Viral agent as cause of disease (B97)	7,162	1.97	78	2.12	319	2.32
Absorbable sulfonamide antibacterial agents	6,718	1.85	68	1.84	124	0.9
Antiemetic/antivertigo agents <sup>a</sup>	6,595	1.81	76	2.06	129	0.94
Antitussives, non-narcotic	5,986	1.64	49	1.33	213	1.55
Tetracyclines	5,164	1.42	63	1.71	116	0.84
Pneumonia (J18)	4,876	1.34	36	0.98	410	2.98
Antihistamines – 1 <sup>st</sup> generation	4,786	1.32	52	1.41	91	0.66
Other general symptoms and signs (R68) <sup>a</sup>	4,293	1.18	48	1.3	103	0.75
Bronchitis (J40)	4,067	1.12	31	0.84	214	1.55
Fever (R50)	3,840	1.06	15	0.41	937	6.8
Otitis media (H66)	3,325	0.91	33	0.9	117	0.85
Gastroenteritis and colitis (K52)	3,147	0.86	35	0.95	89	0.65
Viral infection of site (B34)	2,823	0.78	30	0.81	303	2.2
Analgesic/antipyretics, non-salicylate <sup>a</sup>	2,637	0.72	29	0.79	112	0.81
Respiratory disorder (J98)	2,209	0.61	22	0.6	63	0.46
Influenza (J11)	1,597	0.44	6	0.16	444	3.22
Cephalosporin -3 <sup>rd</sup> generation	1,385	0.38	5	0.14	113	0.82
Cephalosporin -2 <sup>nd</sup> generation	1,321	0.36	13	0.35	45	0.33
Local anesthetics <sup>a</sup>	1,294	0.36	9	0.24	77	0.56
Expectorants	992	0.27	9	0.24	35	0.25
Nose preparations, miscellaneous (OTC)	961	0.26	52	1.41	27	0.2
Prostatitis (N41)	730	0.2	5	0.14	29	0.21

Diagnosis (ICD-10 code) or medication	Dropped instances		Low temperature outliers		High temperature outliers	
	N	% <sup>d</sup>		% <sup>d</sup>	N	% <sup>d</sup>
Non-narcotic antitussive and expectorant combinations	667	0.18	4	0.11	24	0.17
Decongestant-expectorant combination	629	0.17	1	0.03	27	0.2
Tonsillitis (J03)	535	0.15	6	0.16	38	0.28
Tubulo-interstitial nephritis (N12) <sup>a</sup>	433	0.12	3	0.08	28	0.2
Viral infection of intestine (A08)	424	0.12	7	0.19	19	0.14
Breast inflammation (N61) <sup>a</sup>	397	0.11	3	0.08	20	0.15
Non-narcotic antitussive-1 <sup>st</sup> generation antihistamine combination	371	0.1	4	0.11	21	0.15
Non-narcotic antitussive-1 <sup>st</sup> generation antihistamine -decongestant analgesic combination	365	0.1	1	0.03	20	0.15
Narcotic antitussive-anticholinergic combination	319	0.09	1	0.03	18	0.13
Influenza due to other identified influenza virus (J10)	167	0.05	0	0	40	0.29
Acute pyelonephritis (N10)	158	0.04	1	0.03	16	0.12
Peritonsillar abscess (J36)	63	0.02	1	0.03	12	0.09
Total <sup>b</sup>	363,891	100	3,686	100	13,770	100

a Not related to infection either directly or indirectly (clinical determination)

b Type 2 diabetes mellitus was excluded from the analysis dataset by LIMIT, while most other chronic diseases were not. For verification, we compared the mean temperatures (and SD) of encounters with type 2 diabetes mellitus (E11), with hypertension (I10) and with hyperlipidemia (E78), chronic conditions that were not excluded by LIMIT (see eTable3). The mean temperature in encounters with type 2 diabetes mellitus (36.55°C (0.39)) was statistically significantly lower than that with hypertension (36.58°C (0.38)) or with hyperlipidemia (36.58°C (0.37)) (both comparisons:  $p < 0.0001$ ) while those with hypertension and with hyperlipidemia were similar ( $p = 0.57$ ).

A mean temperature (and SD) from the analysis set plus type 2 diabetes mellitus was 36.6325°C (0.3511), which is statistically significantly lower than the mean temperature from the analysis sample only ( $p < 0.0001$ ).

c An encounter could include more than one primary ICD-10 code and/or medication

d Column percent

\* The first three characters of the ICD-10 code define the disease category, and the subsequent digits specify such details as disease etiology, severity, and anatomic site. For example, E11 is type 2 diabetes: E11.3 is type 2 diabetes with ocular involvement, E11.31 is unspecified retinopathy, E11.311 is the same with macular edema and E11.32 is diabetes with non-proliferative retinopathy. These distinctions within diabetes would have little impact on oral temperature but, by reducing sample sizes due to their subset, might have a big impact on LIMIT. As another example, A15 is tuberculosis of the respiratory tract, A15.0 is of the lung and A15.9 is "respiratory, unspecified". We believe this enhanced specificity would have little effect on LIMIT. On the other hand, A17 is tuberculosis of the nervous system, very different from A15 — and they would be seen as different by LIMIT.

Physicians who attach the ICD-10 code to the encounter do not always accurately apply sub-categories, leading to lower sensitivity and possibly lower specificity as well (PMID: 35213502). Given these problems, we focus here on the major diseases rather than the minor ICD-10 code distinctions.

(Note that ICD-10 codes were implemented on October 15, 2015, with ICD-9 codes used prior to this date.)

**eTable 2.** Encounters That Did Not Meet Initial Inclusion Criteria Into the Pre-LIMIT Sample

<b>Variable</b>	<b>Exclusion criterion</b>	<b>N</b>
Temperature	< 34°C (93.2°F)	239
	> 104°C (104°F)	22
Age	< 20 years	33,420
	> 80 years*	52,367
Sex	Unknown	39
Height	< 1.37 m (4'6")	12,090
	> 2.13 m (6'10")	1,011
Weight	< 27.33 kg (60.0 lbs)	0
	> 181.44 (400 lbs)	93,749
Height and weight	Both missing	9,880
	Extreme** or both missing	104,622
BMI	0-10	12,708
	>50	5,328
Primary diagnoses and medications	Missing	110
Primary diagnoses	All non-primary	6,379
<b>AT LEAST ONE EXCLUSION CRITERION</b>		<b>105,893</b>

BMI: Body mass index

\* We expected that, compared to people aged 80 years or younger, people over age 80 years would be fewer (generating wider confidence intervals around temperature estimates) and also less healthy (thus less likely to have a "normal" temperature in the first place).

\*\* (Height < 1.37 m or > 2.13 m) or (weight < 27.33 kg or > 181.44 kg)

**eTable 3.** Most Common *ICD-10* Codes ( $\geq 0.3\%$ ) in the Analysis Set, by Frequency

Diagnosis	Code	Frequency	Percent <sup>a</sup>
Encounter for general adult medical examination without abnormal findings	Z00.00	47,281	11.28
Essential (primary) hypertension	I10	28,719	6.85
Encounter for immunization	Z23	7,462	1.78
Encounter for general adult medical examination	Z00.0	6,878	1.64
Hyperlipidemia, unspecified	E78.5	6,381	1.52
Low back pain	M54.5	4,232	1.01
Generalized anxiety disorder	F41.1	4,050	0.97
Unspecified abdominal pain	R10.9	3,984	0.95
Hypothyroidism, unspecified	E03.9	3,502	0.84
Major depressive disorder, single episode, unspecified	F32.9	3,150	0.75
Pure hypercholesterolemia, unspecified	E78.00	3,046	0.73
Chest pain, unspecified	R07.9	2,691	0.64
Error code (too many meanings)	IMO0002	2,656	0.63
Dorsalgia, unspecified	M54.9	2,628	0.63
Pain in unspecified limb	M79.609	2,625	0.63
Gastro-esophageal reflux disease without esophagitis	K21.9	2,586	0.62
Cervicalgia	M54.2	2,515	0.60
Other malaise	R53.81	2,509	0.60
Headache	R51	2,390	0.57
Pain in unspecified knee	M25.569	2,386	0.57
Anxiety disorder, unspecified	F41.9	2,200	0.52
Dizziness and giddiness	R42	2,180	0.52
Rash and other nonspecific skin eruption	R21	2,139	0.51
Unspecified contact dermatitis, unspecified cause	L25.9	2,130	0.51
Dermatitis, unspecified	L30.9	2,099	0.50
Pain in right knee	M25.561	1,919	0.46
Allergic rhinitis, unspecified	J30.9	1,891	0.45
Encounter for other preprocedural examination	Z01.818	1,891	0.45
Insomnia, unspecified	G47.00	1,745	0.42
Prediabetes	R73.03	1,663	0.40
Other abnormal glucose	R73.09	1,607	0.38
Diarrhea, unspecified	R19.7	1,599	0.38
Pain in left knee	M25.562	1,555	0.43
Elevated blood-pressure reading, without diagnosis of hypertension	R03.0	1,538	0.37

<b>Diagnosis</b>	<b>Code</b>	<b>Frequency</b>	<b>Percent <sup>a</sup></b>
Encounter for gynecological examination (general) (routine) without abnormal findings	Z01.419	1,531	0.37
Epigastric pain	R10.13	1,423	0.34
Palpitations	R00.2	1,414	0.34
Pain in unspecified shoulder	M25.519	1,410	0.34
Other chest pain	R07.89	1,390	0.33
Other fatigue	R53.83	1,376	0.33
Dysuria	R30.0	1,364	0.33
Disorder of the skin and subcutaneous tissue, unspecified	L98.9	1,283	0.31
Frequency of micturition	R35.0	1,240	0.30

a Total diagnoses/medications: N=419,131; total cumulative percent in table: 43.01%.

**eTable 4.** Linear Regression and Mixed-Effects Modeling of Temperature as a Function of Time of Day, Age, Height, Weight and Month, in Men and Women

Parameters	Linear <sup>a,b</sup>		Mixed effects <sup>a,c</sup>	
	Estimate	SE	Estimate	SE
<b>Men (156,826 encounters)</b>				
Intercept	36.154	0.029	36.231	0.035
Hour of day	0.087	0.003	0.075	0.003
Hour of day <sup>2</sup>	-0.003	0.0001	-0.002	0.0001
Age (years)	0.003	0.0004	0.002	0.0005
Age <sup>2</sup>	-0.00005	0.000004	-0.00004	0.000004
Height (m)	-0.174	0.012	-0.167	0.015
Weight (kg)	0.001	0.00006	0.001	0.00007
Month	0.007	0.001	0.008	0.001
Month <sup>2</sup>	-0.0005	0.00008	-0.0006	0.00008
<b>Women (239,370 encounters)</b>				
Intercept	36.256	0.023	36.316	0.027
Hour of day	0.082	0.003	0.077	0.003
Hour of day <sup>2</sup>	-0.003	0.0001	-0.002	0.0001
Age (years)	0.002	0.0003	0.002	0.0004
Age <sup>2</sup>	-0.00005	0.000003	-0.00005	0.000004
Height (m)	-0.111	0.010	-0.132	0.013
Weight (kg)	0.0006	0.00004	0.0008	0.00006
Month	0.009	0.0009	0.009	0.0008
Month <sup>2</sup>	-0.0007	0.00007	-0.0007	0.00006

SE: standard error.

<sup>a</sup> P-values for all estimates: < 0.0001

<sup>b</sup> R<sup>2</sup>: 0.052 (men); 0.055 (women)

<sup>c</sup> AIC: 92,902 (men); 134,067 (women)

## **eAppendix. Missing Data—Weight and Height**

All encounters in the analysis dataset included a BMI measurement. Of these 396,196 encounters, a total of 279,840 (95%) had weight measurements, while 20,169 (5.09%) did not. Among this group missing weight values: 2,213 (11%) were single (i.e. one-time) encounters and as such could not be allocated an estimated weight; 5,438 (27%) had another encounter within 30 days (before or after) that could provide an approximate weight measurement and 12,518 (62%) had another visit beyond 30 days, so that an estimated weight measurement might be considered less reliable.

Overall mean temperature excluding (a) none of these encounters missing a weight, (b) all of these encounters, or (c) all of the encounters either without another visit or with a visit beyond 30 days were similar (respectively: (a) 36.64°C (standard deviation (SD): 0.35°C) [N=396,196], (b) 36.64°C (SD: 0.35°C) [N=376,027], and (c) 36.64°C (SD: 0.35°C) [N=381,465]). Thus, the effect of weight imputation appears to be minimal on our primary variable of interest. Slight differences, however, do become visible beyond the 2 digits after the decimal point, possibly due to the non-randomness of missing data (see below).

We found that weight values were not missing at random. Among patients with 1 encounter, 2-5 encounters and 6 or more encounters, 1.3%, 4.1% and 5.2%, respectively, were missing weight values; the proportion of women increased across these categories. Additionally, patients missing height or weight information were older than those not missing these data. Thus, older women with multiple visits were more likely than others to be missing height and weight information. Reasons for this bias is entirely speculative but may be due to being seen more often and/or regularly. However, it is still worth noting that missing data are comparatively infrequent and have little to no impact on the overall mean values for temperature.

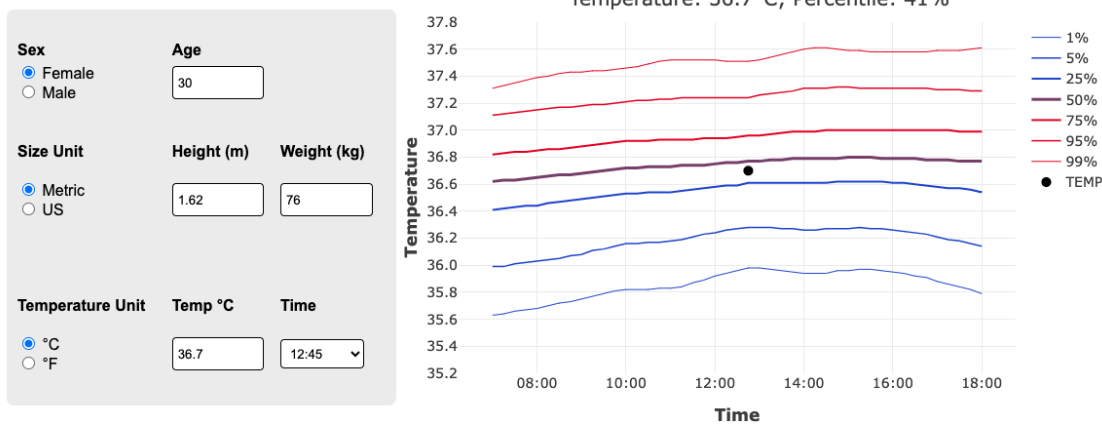
Whereas weight can vary significantly across time, both up and down, changes in adult height between encounters are likely minimal. In fact, adult height is usually not measured at most clinical visits. The missing height value can be estimated from the encounter's BMI value, else the closest next visit.



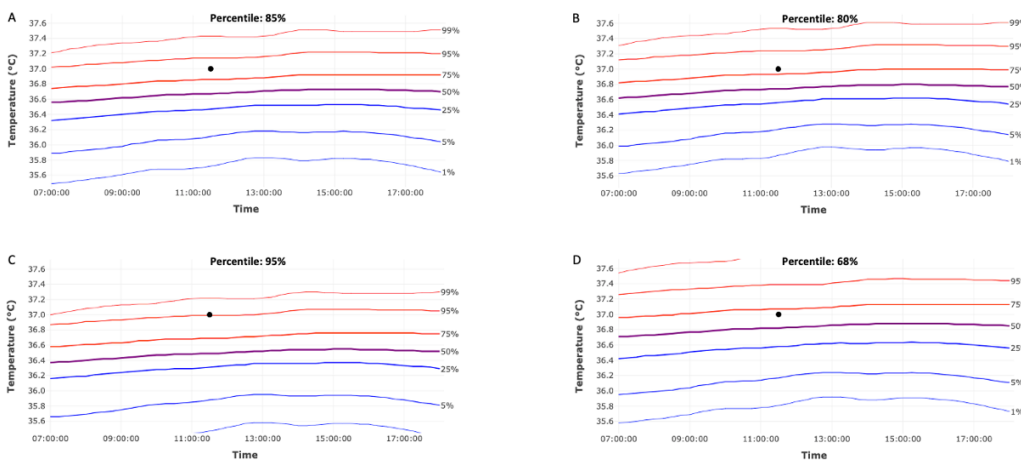
## eFigure. View of the Interactive Visualization Tool for Individualized Normal Temperature Ranges With Examples

An interactive visualization tool that generates individualized normal temperature ranges is available at: <http://normaltemperature.stanford.edu>. (1) The user provides sex, age, height, weight and a temperature measurement with time of day taken. (2) On the right side of the screen, a graph of the normal temperature ranges across the day is generated for this set of individual and temporal characteristics. The temperature measurement is placed on the graph as a black dot. The percentile line indicates where the dot lies within the range of normal temperature; the percentile value indicates the percent of normal temperature values that lie below the dot. If the dot is placed below the 1% line or above the 99% line then this temperature value falls outside of the normal range for this set of characteristics.

In the example below, a 30 year-old woman of height 1.62 m and of weight 76 kg has a temperature of 36.7°C at 12:45 hours. This temperature value (the black dot) falls in the approximate middle of the normal temperature range for this person (41% of “normal” temperatures are below this value).



The 4 graphs below present examples of normal temperature ranges over the day for (A) the average US man (aged 30 years, height 1.73m, weight 88.8 Kg); (B) the average US woman (aged 30 years, height 1.62m, weight 76.4Kg); (C) a man aged 79 years, height 1.83m, weight 68.0Kg; (D) a woman aged 20 years, height 1.40m, weight 150.0Kg. Each individual has the same oral temperature of 37.0°C observed at 11:50 hours (the black dot). The percentile of normal temperature for this single value, printed above each graph, can be seen to vary substantially according to the characteristics of each individual. For the young obese woman, 68% of her “normal” values fall below 37.0°C, whereas for the old thin man, 95% of his expected normal temperature measures fall below this same value.



## **eMethods. Data and Code (SAS and R)**

**Both data and code are available at DRYAD:**

<https://doi.org/10.5061/dryad.3tx95x6mt>

### **1. DATA**

#### **A. Title of dataset:**

Anon\_AnalysisDataV2.csv

#### **B. Summary**

From the Stanford Research Repository (STARR, also known as STRIDE), we identified all adult outpatient encounters that included temperature measurements from April 2008 through June 2017 in the Divisions of Internal Medicine and Family Medicine within Stanford Health Care (Stanford, CA). Oral temperature, the date and time of the temperature measurement, age, sex, weight, height, body mass index (BMI), primary reason for the visit, prescribed medications, and all visit ICD-10 codes were identified from each encounter. An individual patient could have multiple encounters.

After identifying ineligible encounters in STARR, we applied a filtering algorithm, LIMIT, which iteratively removed encounters with primary diagnoses over-represented in the tails of the temperature distribution, leaving only those diagnoses unrelated to temperature. All encounters in the dataset fall into one of these three categories.

The three categories or subsets are specified using the variable "dropped":

- (1) "Baseline" or STARR/STRIDE encounters only (those excluded from consideration by LIMIT as ineligible, due to extreme or missing variables, or highly skewed values (temperatures less than 34°C (93.2°F) or over 40°C (104°F), age less than 20 or over 80 years, BMI less than 10 or over 50, height less than 1.37m (4'6") or greater than 2.13m (6'10"), weight less than 27.22 kg (60.0lbs) or greater than 181.44 kg (400lbs)). [dropped = 99]
- (2) preLIMIT: the "outlier" set of encounters removed by LIMIT due to the identified association between temperature and primary diagnoses. [dropped = 1]
- (3) postLIMIT: The "analysis" set of encounters with usual or "normal" temperature values as determined by LIMIT [dropped = 0].

#### **C. Description of the data and file structure**

All data are found in a single compressed CSV file.

This file is modified from the original analysis data file due to anonymity concerns (see below). Empty cells are intentional, representing no data in the modified set.

#### **D. Sharing/Access information**

This research used data provided by STARR, "STANford medicine Research data Repository," a clinical data warehouse containing live Epic data from Stanford Health Care, the Stanford Children's Hospital, the University Healthcare Alliance and Packard Children's Health Alliance clinics and other auxiliary data from Hospital applications such as radiology PACS. STARR platform is developed and operated by the Stanford Medicine Research IT team and is made possible by the Stanford School of Medicine Research Office.

Our original dataset includes one direct identifier (date/timestamp) and several indirect identifiers: gender, exact age (in SAS-time and normal years), exact weight (metric and US), exact height (metric and US), ethnicity and the department type. Additionally, all encounters were within the same geographic region, considered an additional identifier.

To anonymize the data, we have:

- (1) Removed from each record in the dataset:
  - Date/time stamp

- Year
- BMI
- Ethnicity
- Department type (Family medicine, internal medicine)
- Age in SAS time
- Weight in lbs
- Height in feet, in

(2) Modified within each record:

- Weight in kg by up to +/- 111 g by adding a random number between -0.1111 to 0.1111 to each encounter's weight value.
- Height in meters by up to +/- 2 cm by adding a random number between -0.02 to 0.02 to each encounter's height value.

We have retained the following 3 indirect identifiers;

- Gender (male/female/missing)
- Age (20-80 years)
- Overall geographic region

NOTES:

- The two measures of time that remain are (1) hour of day and (2) month, assessed over a 10-year period.
- Characteristics for encounters where "dropped" = 99 are empty due to the anonymization.
- For the preLIMIT and postLIMIT sets, the change in range by jittering is as follows for: (1) weight: original: 30.38-181.40 kg vs. jittered: 30.29-181.33 kg; (2) height: original: 1.38-2.13 m vs. jittered: 1.36-2.15 m.
- In the case that more detailed information is sought, please contact us to set up a data sharing agreement.

### E. Data Dictionary

Variable	Numeric or character	Label	Comment
age_yrs	Num	Age in yrs	
anon_id	Num	Person ID	
DX1-42	Char	Diag	ICD-10 code for DIAG1 through DIAG42
dropped	Num	99:STRIDE 1:preLIMIT 0:postLIMIT	* Specifies the 3 datasets as listed above
gender	Char	Male/Female	
height_m	Num	Height, m	Jittered (up to +- 2 cm)
Included	Num	Analysis set	
MED1-73	Char		* ICD-10 code for MED1 through MED42
MED_CLASS1-73	Char		* Medication class for MED_CLASS1 through MED_CLASS42
NEW_MED1-73	Char		
primary1-42	Char		Primary diag 1-42 1:Y
primary_dx	Char		Single primary diagnosis for visit
temp_c	Num	Temp in C	
time_hour	Num	Visit: hour of day	
time_month	Num	Visit: month	
weight_kg	Num	Weight, kg	Jittered (up to +- 0.111 kg)

## 2. CODE

Code for this study was written in SAS and/or R:

1. Pre-pre-LIMIT [N=724,199] (pre-preLIMIT only: dropped = 99; Baseline)
  - a. Characteristics of only pre-preLIMIT encounters [eTable 2]  
[Section A: SAS]
2. Pre-LIMIT [N=618,306] (preLIMIT only: dropped = 1);
  - a. LIMIT algorithm  
[Section C: Limit.Rmd]
  - b. Diagnoses removed from dataset (i.e., those associated with temperature) [eTable 1]  
[Section C: Limit.Rmd]
3. Post-LIMIT [N=396,196] (post-LIMIT only: dropped = 0)
  - a. Diagnoses remaining in analysis set (i.e., those not associated with temperature) [eTable 3]  
[Section A: SAS]
  - b. Descriptive analysis comparing outlier set, analysis set and both combined [Table 1]  
[Section A: SAS; also in part in Section C: LIMIT.Rmd]
  - c. Modeling: linear regression and mixed effects modeling [eTable 4]  
[Section A: SAS; also in Section C: LIMIT.Rmd]
  - d. Predictive modeling and quartiles [app]  
[Section B: R Markdown: Temperature Modeling]  
Note that the dataset BH.csv used in predictive modeling with R was a similar but not identical dataset to "AnalysisData".