



SFHERE (Social and Family History -
Extraction, Representation, and Evaluation) Project

Revision History

<i>Version</i>	<i>Description</i>
2014	Guidelines at time of annotation and dissemination <i>Chen ES, Carter EW, Winden TJ, Sarkar IN, Wang Y, Melton GB. Multi-source development of an integrated model for family health history. J Am Med Inform Assoc. 2015 Apr;22(e1):e67-80. doi: 10.1136/amiqnl-2014-003092. PubMed PMID: 25336591.</i>

This document includes the schema and guidelines for annotating **family history** information (Entities, Attributes, and Relationships) in clinical text (e.g., using the Brat Rapid Annotation Tool [BRAT]).

Table of Contents

1. General Guidelines
2. Annotation Guidelines and Examples – Entities and Attributes
3. Annotation Guidelines and Examples – Relationships



1. General Guidelines

- Annotate as much information in the sentence as possible, but do not need to annotate words like “has”, “had”, “is”, etc.
- Use “Notes” field at bottom of “Edit Annotation” window for any notes or comments (this will show a blue box around the annotation)
 - If annotation is not related to family history (e.g., social history information), put the code “NOT FH” in the Notes field
 - If annotation is related to the patient, put the code “PATIENT” in the Notes field
- Attributes should be specified for the following entities (further detailed in Section 2; this will show a “*” in the annotation)
 - ObservationType for “Observation” or “SubObservation”
 - SideFamilyType for “Side of Family”
 - CertaintyDegree for “Certainty of Observation” or “Certainty of Subject”
 - TimeSpecificity for “Current Age”, “Age of Diagnosis”, “Date of Diagnosis”, “Age of Death”, “Date of Death”, or “Temporality of Observation”
- Do not include punctuation (“.”, “,”, etc.) in annotation where possible

2. Annotation Guidelines and Examples – Entities and Attributes

Entity (Variable Name)	Attribute	Description	Examples
SUBJECT INFORMATION			
Person (Person)		Any other person mentioned in the sentence (not patient of family member)	“her <u>grandmother</u> who is present at visit today”
Source of Information (Source)		Person specified as providing information about family history (other than patient)	“ <u>mom</u> unsure of what type of cancer”
Subject (Subject)		If unclear if family member of patient, annotate as Subject and put a note	
Patient (Patient)		If information is related to the patient rather than family member	“patient”; “he”
Family Member (FM)		Family member <ul style="list-style-type: none"> • If both family member and side of family are specified in the sentence, annotate them separately (e.g., ‘paternal aunt’ – “paternal” is side of family and “aunt” is family member). • Do not specify side of family if it is not explicitly stated or is a first degree relative (e.g., mother or father) • Include husband/wife • If “he” or “she” refers to family member 	“mother”; “father”; “child”; “grandmother”; “uncle”; “paternal <u>grandfather</u> ”; “maternal <u>aunt</u> ”; “MA”, “MGM”; “ <u>one</u> of them”; “ <u>two</u> had”



		<p>mentioned in previous sentences, highlight as Family Member and put a note (e.g., refers to father from previous sentence)</p> <ul style="list-style-type: none"> • If abbreviation used (e.g., “MA” for maternal aunt or “MGM” for maternal grandmother”), annotate as Family Member and put a note (e.g., includes side of family) • If there is a quantity or number but no family member specified (e.g., refers to family members earlier in the sentence), annotate the number as Family Member and put a note 	
Side of Family (SideFamily)		Side of family	“ <u>paternal</u> grandfather”; “ <u>maternal</u> aunt”
	<i>Other</i>		
	<i>Maternal</i>		“maternal”; “mother’s side”; “maternal side of family”
	<i>Paternal</i>		“paternal”; “father’s side”; “paternal side of family”
	<i>Both</i>		“both sides of family”
Vital Status (SubjVitalStatus)		Indicates whether family member is alive or deceased	“early death”; “passed away”; “died”; “unknown”
Current Age (SubjCurrentAge)		May be non-specific or specific age; also specify “TimeSpecificity” attribute	“elderly”; “60s”; “age 40”, “late 60s”
	<i>InexactQuantitative</i>		“60s”; “late 60s”
	<i>Other</i>		
	<i>Exact</i>	Highlight “age of”, “years of age”, etc. as part of annotation but do not include non-age/date related words such as “ <u>in her</u> 80s” or “ <u>her</u> 80s”	“age 40”; “age of 83”; “32 years of age”
	<i>InexactQualitative</i>		“elderly”
Race or Ethnicity (SubjRaceEthnicity)			“Caucasian”; “Hispanic”
Negation of Subject (SubjNegation)		Negates the subject/family member	“No siblings”
SubjCertainty (SubjCetainty)		Qualifies the certainty of information about the subject	“Possible uncle or cousin with colon cancer”



	<i>Unknown</i>		
	<i>Other</i>		
	<i>Uncertain</i>		
	<i>Certain</i>		
Quantity of Subject (SubjQuantity)		Number of family members	“ <u>one</u> grandchild”; “ <u>4</u> sisters”; “ <u>two</u> family members”
Other Modifier of Subject (SubjOtherMod)		Any other modifier not covered by certainty, negation, etc.	
OBSERVATION INFORMATION			
Observation (Observation)		Disease, procedure, or other historical observation element <ul style="list-style-type: none"> Issue: cannot use fragments in multiple annotations – see SubObservation below for how to handle cases such as “alcohol or tobacco use” A statement may include several related observations – if possible, they should be annotated separately but the linkage will be lost 	“diabetes”; “breast cancer”; “hypertension”; “ <u>blindness secondary to diabetic retinopathy</u> ”; “ <u>pneumonia as a complication to Alzheimer’s disease</u> ”
	<i>LabTest</i>	Describes a lab test or result of a test	“BRCA neg”
	<i>Other</i>		“healthy”; “ill”; “well”; “vegetarian”
	<i>Disease</i>	Describes a disease or condition	“hypothyroidism”; “heart disease”; “prematurely born”; “tobacco dependence”; “old age”; “non-smoker”; “smoker”; “ex-smoker”
	<i>Procedure</i>	Describes a medical procedure or treatment	“appendectomy”; “lumpectomy”
	<i>Medication</i>	Describes a medication	“on medication to lower it” (in this case, put note about how this relates to elevated cholesterol mentioned earlier in the sentence)
SubObservation		For cases such as “x, y, z condition” <ul style="list-style-type: none"> “condition” should be annotated as “Observation” and “x”, “y”, and “z” as separate “SubObservation” Each “SubObservation” should be linked to the “Observation” Any modifiers that apply to “Observation” should be linked to the respective “SubObservation” instead 	“current breast, previous colon, and definite ovarian cancer” Observation = cancer SubObservation = breast ObsTemporal = current SubObservation = colon ObsTemporal = previous SubObservation = ovarian



			ObsCertainty = definite “alcohol or tobacco use” Observation = use SubObservation = alcohol SubObservation = tobacco
Negation of Observation (ObsNegation)		Negates the observation	“ <u>No</u> paternal history of colon cancer”; “Sister <u>without</u> heart disease or stroke”; “ <u>no</u> family history of bleeding disorder”
Certainty of Observation (ObsCertainty)		Qualifies the certainty of the observation	“There is”; “positive”
	<i>Unknown</i>		
	<i>Other</i>		
	<i>Uncertain</i>		“Mother with <u>probable</u> blood clot”; “Likely”
	<i>Certain</i>		“there is”; “positive”
Quantity of Observation (ObsQuantity)		Quantity of observation	“ <u>multiple</u> strokes”; “ <u>some</u> issues”
Strength of Observation (ObsStrength)		Strength of the observation	“significant”; “remarkable”; “strong”; “strongly”
Relevance of Observation (ObsRelevance)		Relevance of the observation	“pertinent”; “applicable”; “related”;
Temporality of Observation (ObsTemporal)		For temporal information related to observation	“currently”; “past”; “previous”; “remote” DO NOT annotate “history of” or “family history of”
Other Modifier of Observation (ObsOtherMod)		Any other modifier not covered by certainty, negation, strength, etc.	“blindness <u>secondary to</u> diabetic retinopathy”; “pneumonia as a <u>complication to</u> Alzheimer’s disease”
AGE/DATE INFORMATION			
Age of Diagnosis (AgeDiag)	See Current Age for TimeSpecificity Attributes	May be non-specific or specific age; also specify “TimeSpecificity” attribute	“elderly”; “60s”; “age 40”; “late 60s”
Date of Diagnosis (DateDiag)		May be non-specific or specific date; also specify “TimeSpecificity” attribute	“January 1, 1965”; “1960”; “around 1980”; “recently”
	<i>InexactQuantitative</i>		“around 1980”; “1990-1995”
	<i>Other</i>		



	<i>Exact</i>		“January 1, 1965”; “1960”; “May 1970”;
	<i>InexactQualitative</i>		“awhile ago”; “recently”; “current”; “many years ago”
Age of Death (AgeDeath)	See Current Age for TimeSpecificity Attributes	May be non-specific or specific age; also specify “TimeSpecificity” attribute	“elderly”; “60s”; “age 40”; “late 60s”
Date of Death (DateDeath)	See Date of Diagnosis for TimeSpecificity Attributes	May be non-specific or specific date; also specify “TimeSpecificity” attribute	“January 1, 1965”; “1960”; “around 1980”; “recently”
OTHER INFORMATION			
OtherModifier (OtherModifier)		Any other modifier not covered by defined modifiers for subject and observation	
Other Family History Text (OtherStatement)		No observation and no subject, but is related to family history. <ul style="list-style-type: none"> • Can also mark related text. • Can also add a note at the bottom. 	“noncontributory to this illness”; “no other information is provided”; “noncontributory to the present condition”; “reviewed and unremarkable”; “of Korean origin”; “noncontributory”; “unremarkable”; “noted and reviewed today”; “reviewed and unchanged”

3. Annotation Guidelines and Examples – Relationships

Relation Label	Description	From (Arg1)	To (Arg2)	Examples
Relationships for Subject, Family Member, or Side of Family				
SideAttrSubj	Side of family attributed to subject	Side of family	Subject or FM or Patient	Maternal ⇒ grandmother
VitalStatusSubj	Vital status of subject	Subject vital status	Subject or FM or Patient	Passing away ⇒ mother
CurrentAgeSubj	Current age of subject	Subject current age	Subject or FM or Patient	Age 45 ⇒ brother
RaceEthnicitySubj	Race/ethnicity of subject	Subject race/ethnicity	Subject or FM or Patient	Caucasian ⇒ father
QuanSubj	Quantity of subject	Subject quantity	Subject or FM or Patient	4 ⇒ sisters
OtherModSubj	Other modifier of subject	Subject other mod	Subject or FM or Patient	
NegSubj	Negation of subject	Subject negation	Subject or FM or Patient	None ⇒ uncle
CertainSubj	Certainty of subject	Subject certainty	Subject or FM or Patient	
CertainSide	Certainty of side of family	Subject certainty	Side of family	
ObsFamDeath	FM death observation	Observation	Subject or FM or Patient	Colon cancer ⇒ father
AgeDiagFam	Age of diagnosis in FM	Age at diagnosis	Subject or FM or side of family or Patient	At age 32 ⇒ sister
DateDiagFam	Date of diagnosis in FM	Date of diagnosis	Subject or FM or side of family or Patient	Late 60s ⇒ mother



SFHERE (Social and Family History -
Extraction, Representation, and Evaluation) Project

Annotation Guidelines for
Family History

AgeDeathFam	Age of death of FM	Age of death	Subject or FM or side of family or Patient	Late 70s ⇒ mother
DateDeathFam	Date of death of FM	Date of death	Subject or FM or side of family or Patient	2001 ⇒ grandfather
Relationships for Observation				
FamObsArg	FM observation	Subject or FM	Observation	mother ⇒ diabetes
SideObsArg	Side of family observation	Side of Family	Observation	Paternal ⇒ heart disease
AgeDiagObs	Age of diagnosis of observation	Age diagnosed	Observation	Elderly ⇒ hypertension
DateDiagObs	Date of diagnosis	Dated diagnosed	Observation	1980 ⇒ breast cancer
NegObs	Negation of observation	Observation negation	Observation	No ⇒ bleeding disorder
CertainObs	Certainty of observation	Observation certainty	Observation	Positive ⇒ heart attack
QuantityObs	Quantity of observation	Observation quantity	Observation	Multiple ⇒ strokes
StrengthObs	Strength of observation	Observation strength	Observation	Strong ⇒ hypertension
RelevanceObs	Relevance of observation	Observation relevance	Observation	Pertinent ⇒ hypertension
TemporalObs	Temporality of observation	ObsTemporal	Observation	Currently ⇒ melanoma
OtherModObs	Other	Observation other mod	Observation	
Relationships for SubObservation				
SubObs	Link subobservation to observation	SubObservation	Observation	colon ⇒ cancer
FamSubObsArg	FM subobservation	Subject or FM	SubObservation	mother ⇒ colon
SideSubObsArg	Side of family subobservation	Side of Family	SubObservation	Paternal ⇒ colon
AgeDiagSubObs	Age of diagnosis of subobservation	Age diagnosed	SubObservation	Elderly ⇒ colon
DateDiagSubObs	Date of diagnosis of subobs	Dated diagnosed	SubObservation	1980 ⇒ colon
NegSubObs	Negation of subobservation	Observation negation	SubObservation	No ⇒ colon
CertainSubObs	Certainty of subobservation	Observation certainty	SubObservation	Positive ⇒ colon
StrengthSubObs	Strength of subobservation	Observation strength	SubObservation	Strong ⇒ colon
RelevanceSubObs	Relevance of subobservation	Observation relevance	SubObservation	Pertinent ⇒ colon
TemporalSubObs	Temporality of subobservation	ObsTemporal	SubObservation	Currently ⇒ colon
OtherModSubObs	Other	Observation other mod	SubObservation	