

The context of application of these Statistic Model Codes will be:

Structure

Name	Flags	Card.	Type	Description & Constraints
Statistic	Σ D		Element	Single statistic Elements defined in Ancestors: <i>id</i> , <i>extension</i> , <i>modifierExtension</i>
description	Σ	0..1	string	Description of content
note	Σ	0..*	Annotation	Footnotes and/or explanatory notes
statisticType	Σ	0..1	CodeableConcept	Type of statistic, eg relative risk <i>StatisticType (Extensible)</i>
quantity	Σ	0..1	Quantity	Statistic value
numberOfEvents	Σ	0..1	unsignedInt	The number of events associated with the statistic
sampleSize	Σ	0..1	Element	Number of samples in the statistic
attributeEstimate	Σ	0..*	Element	An attribute of the Statistic
modelCharacteristic	Σ	0..*	Element	Model characteristic
code	Σ	1..1	CodeableConcept	Model specification <i>StatisticModelCode (Extensible)</i>
value	Σ	0..1	SimpleQuantity	Numerical value to complete model specification
variable	Σ	0..*	Element	A variable adjusted for in the adjusted analysis
variableDefinition	Σ	1..1	Reference(Group EvidenceVariable) code	Description of the variable
handling	Σ	0..1	code	continuous dichotomous ordinal polychotomous <i>EvidenceVariableHandling (Required)</i>
valueCategory	Σ	0..*	CodeableConcept	Description for grouping of ordinal or polychotomous variables
valueQuantity	Σ	0..*	Quantity	Discrete value for grouping of ordinal or polychotomous variables
valueRange	Σ	0..*	Range	Range of values for grouping of ordinal or polychotomous variables

The **Statistic Standard and Terminology Working Group** discussed the overall framework for coordinating the Statistic Model Code System with the STATO Statistics Ontology.

We determined the term “statistical model” can be reclassified as an “information content entity” in STATO and we can then map out statistical model terms similar to what we did for “statistic” terms for our Statistic Type Code System.

However, we also determined that some of the terms we want to include in our Statistic Model Code System (to support expression in *Evidence.statistic.modelCharacteristic.code* elements) are technically not statistical models. These terms are likely statistical processes and be found in the “statistical hypothesis test” class within STATO. For these terms we will map terms following the current outline in STATO and see how well it fits our functional application.

- Statistical model
 - Regression model
 - Regression Model Form
 - Linear Regression (GLM with identity link)
 - Logistic Regression (GLM with logit link)
 - Log Linear Regression
 - Polynomial Regression
 - Cox Proportional Hazards

- Regression Model Distribution
 - Normal Distribution for Regression
 - Log Normal Distribution for Regression
 - Exponential Family of Distributions for Regression
 - Binomial Distribution for Regression (Binomial Regression)
 - Multinomial Distribution for Regression (Multinomial Regression)
 - Poisson Regression (Poisson Distribution for Regression)
 - Negative Binomial Regression (Negative Binomial Distribution for Regression)
 - GLM (Generalized Linear Model)
 - GLM with probit link
 - GLM with logit link (Logistic Regression)
 - GLM with identity link (Linear Regression)
 - GLM with log link
 - GLM with generalized logit link
 - GLM with subtype unspecified
 - Generalized Linear Mixed Model (GLMM)
 - GLMM with probit link
 - GLMM with logit link
 - GLMM with identity link
 - GLMM with log link
 - GLMM with generalized logit link
 - GLMM with subtype unspecified
- Data transformation
 - o Statistical hypothesis test
 - Between group comparison statistical test
 - ANOVA
 - o multivariate ANOVA (MANOVA)
 - o Multiway ANOVA
 - ?? 3-way ANOVA
 - o One-way ANOVA
 - o Repeated measure ANOVA
 - o Two-way ANOVA
 - ?? 2-way ANOVA without replication
 - ?? 2-way ANOVA with replication
 - Non-parametric test
 - o Kruskal Wallis test
 - o Log rank test
 - o Mann-Whitney U-test (?? Wilcoxon Rank-Sum test; U test; Wilcoxon rank-sum test; rank-sum test for the comparison of two samples)
 - o McNemar test ((move from information content entity))
 - o Sign test
 - o Friedman test
 - Two sample t-test (2-sample t-test, independent)
 - o Two sample t-test with equal variance
 - o Two sample t-test with unequal variance

- Z test for between group comparison
- Chi square test
 - Pearson's Chi square test of goodness of fit
 - Pearson's Chi square test of goodness of independence between categorical variables
 - o Yate's corrected Chi-Squared test
- Single-sample reference comparison statistical test
 - One sample t-test (1-sample t-test)
 - Z test for single-sample
- Test of association between categorical variables
 - Cochran-Armitage test for trend
 - Fisher's exact test
- Within subject comparison statistical test
 - Paired t-test (2-sample t-test, dependent, matched pair t-test)
 - Wilcoxon signed rank test

In process of moving up --- The current proposed set of concepts with draft terms (for the value set) for the Statistic Model Code System (assuming t-test is added to the [Statistic Model Code System Step 3 spreadsheet](#)) is (with items bolded if they are both a classifier and a codable concept):

1. Determination of Relationship codes (CATEGORY ONLY)
 - a. **Threshold**
 - i. alpha setting
 1. alpha setting with subtype unspecified
 2. individual test alpha without multiple testing adjustment
 3. overall alpha with multiple testing
 4. individual test alpha with multiple testing adjustment
 - ii. one-tailed test (one threshold)
 - iii. two-tailed test (two thresholds)
 - b.
 - c. Threshold framing (CATEGORY ONLY)
 - i. **alpha setting**
 1. overall alpha with multiple testing
 2. individual test alpha with multiple testing adjustment
 - ii. alpha setting (CATEGORY ONLY)
 1. alpha setting without known subtype
 2. individual test alpha without multiple testing adjustment
 3. overall alpha with multiple testing
 4. individual test alpha with multiple testing adjustment
 - iii. one-tailed test (one threshold)
 - iv. two-tailed test (two thresholds)
2. Adjustment of Variables codes (CATEGORY ONLY)
 - a. Adjustment for confounders
 - b. Adjustment for clustering
 - c. Zero-cell adjustment with constant
 - d. Zero-cell adjustment with continuity correction
 - e. Interaction term

- f. Mantel-Haenszel method
- 3. Meta-analysis**
 - a. Inverse variance method
 - b. Peto method
 - c. Adjustment of variance codes (CATEGORY ONLY)
 - i. Hartung-Knapp adjustment
 - ii. Modified Hartung-Knapp adjustment
- 4. Variance codes (CATEGORY ONLY)
 - a. Effect size distribution assumption codes (CATEGORY ONLY)
 - i. Fixed-effect model (common-effect model, one true effect size)
 - ii. Random-effects model (random effects, true effect sizes are distributed)
 - b. Heterogeneity codes (CATEGORY ONLY)
 - i. Chi-square test for homogeneity
 - ii. Tau estimation (CATEGORY ONLY)
 - 1. Dersimonian-Laird method
 - 2. Paule-Mandel method
 - 3. Restricted Maximum Likelihood method
 - 4. Maximum Likelihood method
 - 5. Empirical Bayes method
 - 6. Hunter-Schmidt method
 - 7. Sidik-Jonkman method
 - 8. Hedges method