Endpoint	**Genotoxicity** in vitro comet assay
Endpoint description	Genotoxicity – genotoxic effect in mammalian cell lines treated in vitro with TiO_2 nanoforms, using comet assay performed according to scientifically valid, recognized protocols, including those described in Collins, A., Møller, P., Gajski, G. et al. Measuring DNA modifications with the comet assay: a compendium of protocols. Nat Protoc 18, 929–989 (2023).
Nanoform	TiO_2
Data quality control	The QSAR model utilizes extracted literature data subjected to a quality control including the evaluation of: i) relevance of the test material, ii) reliability of genotoxicity studies, and iii) relevance of genotoxicity studies.
Type of model	L-PCA + kNN
Descriptors	The model includes descriptors of nanoform physicochemical characteristics and cell line characteristics.
Dataset	

	training set	**validation set**	**total**
positive	69	18	111
negative	19	5	

Statistics

	training set	**validation set**
accuracy	0.977	0.870
precision	1.000	0.895
recall	0.971	0.944
F1 score	0.985	0.919
MCC	0.937	0.592

Inclusion criteria to applicability domain

Chemical composition of nanoform: 'TiO_2'
Shape of nanoform: 'spherical'
Crystal structure: 'anatase' OR 'rutile' OR 'anatase+rutile'
Surface area, value up to: 343.7
Minimum particle size, range: 3.9 – 142.0
Mean particle size, range: 3.9 – 280.0
Maximum particle size, range: 4.0 – 400.0

Endpoint	**Genotoxicity** in vitro comet assay
Endpoint description	Genotoxicity – genotoxic effect in mammalian cell lines treated in vitro with SiO_2 nanoforms, using comet assay performed according to scientifically valid, recognized protocols, including those described in Collins, A., Møller, P., Gajski, G. et al. Measuring DNA modifications with the comet assay: a compendium of protocols. Nat Protoc 18, 929–989 (2023).
Nanoform	SiO_2
Data quality control	The QSAR model utilizes extracted literature data subjected to a quality control including the evaluation of: i) relevance of the test material, ii) reliability of genotoxicity studies, and iii) relevance of genotoxicity studies.
Type of model	L-PCA + Decision tree
Descriptors	The model includes descriptors of nanoform physicochemical characteristics and cell line characteristics.
Dataset	

	training set	**validation set**	**total**
positive	34	9	65
negative	18	4	

Statistics

	training set	**validation set**
accuracy	0.865	0.769
precision	0.966	0.800
recall	0.824	0.889
F1 score	0.889	0.842
MCC	0.736	0.426

Inclusion criteria to applicability domain

Chemical composition of nanoform: 'SiO_2'
Shape of nanoform: 'spherical'
Surface area, value up to: 450.0
Minimum particle size, range: 5.0 - 166.1
Mean particle size, range: 6.0 - 169.2
Maximum particle size, range: 6.0 – 172.3