

Supplementary material for the publication:
**Automatic classification of experimental
models in biomedical literature to
support searching for alternative
methods to animal experiments**

Mariana Neves, Antonina Klippert, Fanny Knöspel, Juliane Rudeck,
Ailine Stolz, Zsafia Ban, Markus Becker, Kai Diederich,
Barbara Grune, Pia Kahnau, Nils Ohnesorge, Johannes Pucher,
Gilbert Schönfelder, Bettina Bert and Daniel Butzke

Contents

1	Queries for retrieval of the abstracts	2
2	Annotation guidelines	15
2.1	In vivo experiments on vertebrates, cephalopods	16
2.2	Vertebrate organs and tissues	17
2.3	Vertebrate primary and stem cells	17
2.4	Vertebrate immortalized (transformed) and cancer cell lines . . .	17
2.5	Experiments with humans or human materials	17
2.6	Experiments with invertebrates or invertebrate materials	18
2.7	In silico	18
2.8	Other	18
3	PMIDs removed in the second rounds	20
4	Statistics of the corpus	21
5	Selection of the hyperparameters	24
6	Annotations - error analysis	28
7	Evaluation of the additional rounds	30
8	Analysis of the MeSH terms	31
9	Experiments with section labels	32

1 Queries for retrieval of the abstracts

“In Vivo” ((“Animals, Newborn”[MeSH Terms] OR “Disease Models, Animal”[MeSH Terms] OR “Animals, Genetically Modified”[MeSH Terms] OR “Animal Experimentation”[MeSH Terms] OR “Vivisection”[MeSH Terms] OR “Models, Animal”[MeSH Terms] OR “Xenograft Model Antitumor Assays”[MeSH Terms] OR “Neoplasm Transplantation”[MeSH Terms] OR “Leukemia, Experimental”[MeSH Terms] OR “Liver Neoplasms, Experimental”[MeSH Terms] OR “Mammary Neoplasms, Experimental”[MeSH Terms] OR “Melanoma, Experimental”[MeSH Terms] OR “Sarcoma, Experimental”[MeSH Terms] OR “Encephalomyelitis, Autoimmune, Experimental”[MeSH Terms] OR “Myasthenia Gravis, Autoimmune, Experimental”[MeSH Terms] OR “Neuritis, Autoimmune, Experimental”[MeSH Terms] OR “Arthritis, Experimental”[MeSH Terms] OR “Diabetes Mellitus, Experimental”[MeSH Terms] OR “Liver Cirrhosis, Experimental”[MeSH Terms] OR “Neoplasms, Experimental”[MeSH Terms] OR “Nervous System Autoimmune Disease, Experimental”[MeSH Terms] OR “Radiation Injuries, Experimental”[MeSH Terms] OR “Mice”[MeSH Terms] OR “Mice, Knockout”[MeSH Terms] OR “Mice, Transgenic”[MeSH Terms] OR “Mice, Inbred C57BL”[MeSH Terms] OR “Mice, Nude”[MeSH Terms] OR “Mice, SCID”[MeSH Terms] OR “Mice, Congenic”[MeSH Terms] OR “Mice, Inbred Strains”[MeSH Terms] OR “Mice, Mutant Strains”[MeSH Terms] OR “Mice, Obese”[MeSH Terms] OR “Mice, 129 Strain”[MeSH Terms] OR “Mice, Hairless”[MeSH Terms] OR “Rats”[MeSH Terms] OR “Rats, Sprague-Dawley”[MeSH Terms] OR “Rats, Wistar”[MeSH Terms] OR “Rats, Transgenic”[MeSH Terms] OR “Rats, Mutant Strains”[MeSH Terms] OR “Rats, Long-Evans”[MeSH Terms] OR “Rats, Inbred Strains”[MeSH Terms]) AND “Diseases Category”[MeSH Terms]) NOT (“Clinical Study”[Publication Type] OR “Clinical Trial”[Publication Type] OR “Clinical Trial, Phase I”[Publication Type] OR “Clinical Trial, Phase II”[Publication Type] OR “Clinical Trial, Phase III”[Publication Type] OR “Clinical Trial, Phase IV”[Publication Type] OR “Clinical Trial, Veterinary”[Publication Type] OR “Clinical Trial Protocol”[Publication Type] OR “Controlled Clinical Trial”[Publication Type] OR “Clinical Conference”[Publication Type] OR “Review”[Publication Type]) NOT (“Organ Culture Techniques”[MeSH Terms] OR “Tissue Culture Techniques”[MeSH Terms] OR “Tissue Extracts”[MeSH Terms] OR “Actihaemyl”[MeSH Terms] OR “Cell Extracts”[MeSH Terms] OR “Liver Extracts”[MeSH Terms] OR “Pancreatic Extracts”[MeSH Terms] OR “Pancreatin”[MeSH Terms] OR “Pancrelipase”[MeSH Terms] OR “Placental Extracts”[MeSH Terms] OR “Thymus Extracts”[MeSH Terms] OR “Isolated Heart Preparation”[MeSH Terms]) NOT (“Adult Germline Stem Cells”[MeSH Terms] OR “Adult Stem Cells”[MeSH Terms] OR “Blastomeres”[MeSH Terms] OR “Cellular Reprogramming Techniques”[MeSH Terms] OR “Embryoid Bodies”[MeSH Terms] OR “Embryonal Carcinoma Stem Cells”[MeSH Terms] OR “Embryonal Carcinoma Stem Cells”[MeSH Terms] OR “Embryonic Germ Cells”[MeSH Terms] OR “Embryonic Stem Cells”[MeSH Terms] OR “Fetal Stem Cells”[MeSH Terms] OR “Granulocyte-Macrophage Progenitor Cells”[MeSH Terms] OR “Hematopoietic Stem Cells”[MeSH Terms] OR “Human Embryonic Stem Cells”[MeSH Terms] OR “Induced Pluripotent

Stem Cells"[MeSH Terms] OR "Lymphoid Progenitor Cells"[MeSH Terms] OR "Megakaryocyte-Erythroid Progenitor Cells"[MeSH Terms] OR "Mesenchymal Stem Cells"[MeSH Terms] OR "Mouse Embryonic Stem Cells"[MeSH Terms] OR "Multipotent Stem Cells"[MeSH Terms] OR "Myeloid Progenitor Cells"[MeSH Terms] OR "Myoblasts, Cardiac"[MeSH Terms] OR "Myoblasts, Skeletal"[MeSH Terms] OR "Myoblasts, Smooth Muscle"[MeSH Terms] OR "Myoblasts"[MeSH Terms] OR "Neoplastic Stem Cells"[MeSH Terms] OR "Neural Stem Cells"[MeSH Terms] OR "Oligodendrocyte Precursor Cells"[MeSH Terms] OR "Oogonial Stem Cells"[MeSH Terms] OR "Peripheral Blood Stem Cells"[MeSH Terms] OR "Pluripotent Stem Cells"[MeSH Terms] OR "Precursor Cells, B-Lymphoid"[MeSH Terms] OR "Precursor Cells, T-Lymphoid"[MeSH Terms] OR "Primary Cell Culture"[MeSH Terms] OR "Satellite Cells, Skeletal Muscle"[MeSH Terms] OR "Side-Population Cells"[MeSH Terms] OR "Stem Cell Research"[MeSH Terms] OR "Stem Cells"[MeSH Terms] OR "Thymocytes"[MeSH Terms] OR "Totipotent Stem Cells"[MeSH Terms]) NOT ("3T3 Cells"[MeSH Terms] OR "A549 Cells"[MeSH Terms] OR "BALB 3T3 Cells"[MeSH Terms] OR "Caco-2 Cells"[MeSH Terms] OR "Cell Line, Transformed"[MeSH Terms] OR "Cell Line, Tumor"[MeSH Terms] OR "Cell Line"[All Fields] OR "CHO Cells"[MeSH Terms] OR "COS Cells"[MeSH Terms] OR "HCT116 Cells"[MeSH Terms] OR "HEK293 Cells"[MeSH Terms] OR "HeLa Cells"[MeSH Terms] OR "Hep G2 Cells"[MeSH Terms] OR "HL-60 Cells"[MeSH Terms] OR "HT29 Cells"[MeSH Terms] OR "Jurkat Cells"[MeSH Terms] OR "K562 Cells"[MeSH Terms] OR "L Cells (Cell Line)"[MeSH Terms] OR "LLC-PK1 Cells"[MeSH Terms] OR "Madin Darby Canine Kidney Cells"[MeSH Terms] OR "MCF-7 Cells"[MeSH Terms] OR "NIH 3T3 Cells"[MeSH Terms] OR "PC12 Cells"[MeSH Terms] OR "PC-3 Cells"[MeSH Terms] OR "RAW 264.7 Cells"[MeSH Terms] OR "Sf9 Cells"[MeSH Terms] OR "Swiss 3T3 Cells"[MeSH Terms] OR "THP-1 Cells"[MeSH Terms] OR "Tumor Cells, Cultured"[MeSH Terms] OR "Cells, Cultured"[MeSH Terms] OR "U937 Cells"[MeSH Terms] OR "Vero Cells"[MeSH Terms]) NOT "Computer Simulation"[MeSH Terms] NOT "Humans"[MeSH Terms:noexp] AND (hasabstract[text] AND "Animals"[MeSH Terms:noexp] AND English[lang])

"Organs/Tissues" (("Organ Culture Techniques"[MeSH Terms] OR "Tissue Culture Techniques"[MeSH Terms] OR "Tissue Extracts"[MeSH Terms] OR "Actihaemyl"[MeSH Terms] OR "Cell Extracts"[MeSH Terms] OR "Liver Extracts"[MeSH Terms] OR "Pancreatic Extracts"[MeSH Terms] OR "Pancreatin"[MeSH Terms] OR "Pancrelipase"[MeSH Terms] OR "Placental Extracts"[MeSH Terms] OR "Thymus Extracts"[MeSH Terms] OR "Isolated Heart Preparation"[MeSH Terms]) AND "Diseases Category"[MeSH Terms]) NOT ("Clinical Study"[Publication Type] OR "Clinical Trial"[Publication Type] OR "Clinical Trial, Phase I"[Publication Type] OR "Clinical Trial, Phase II"[Publication Type] OR "Clinical Trial, Phase III"[Publication Type] OR "Clinical Trial, Phase IV"[Publication Type] OR "Clinical Trial, Veterinary"[Publication Type] OR "Clinical Trial Protocol"[Publication Type] OR "Controlled Clinical Trial"[Publication Type] OR "Clinical Conference"[Publication Type] OR "Review"[Publication Type])

Type]) NOT ("Animals, Newborn"[MeSH Terms] OR "Disease Models, Animal"[MeSH Terms] OR "Animals, Genetically Modified"[MeSH Terms] OR "Animal Experimentation"[MeSH Terms] OR "Vivisection"[MeSH Terms] OR "Models, Animal"[MeSH Terms] OR "Xenograft Model Antitumor Assays"[MeSH Terms] OR "Neoplasm Transplantation"[MeSH Terms] OR "Leukemia, Experimental"[MeSH Terms] OR "Liver Neoplasms, Experimental"[MeSH Terms] OR "Mammary Neoplasms, Experimental"[MeSH Terms] OR "Melanoma, Experimental"[MeSH Terms] OR "Sarcoma, Experimental"[MeSH Terms] OR "Encephalomyelitis, Autoimmune, Experimental"[MeSH Terms] OR "Myasthenia Gravis, Autoimmune, Experimental"[MeSH Terms] OR "Neuritis, Autoimmune, Experimental"[MeSH Terms] OR "Arthritis, Experimental"[MeSH Terms] OR "Diabetes Mellitus, Experimental"[MeSH Terms] OR "Liver Cirrhosis, Experimental"[MeSH Terms] OR "Neoplasms, Experimental"[MeSH Terms] OR "Nervous System Autoimmune Disease, Experimental"[MeSH Terms] OR "Radiation Injuries, Experimental"[MeSH Terms] OR "Mice"[MeSH Terms] OR "Mice, Knockout"[MeSH Terms] OR "Mice, Transgenic"[MeSH Terms] OR "Mice, Inbred C57BL"[MeSH Terms] OR "Mice, Nude"[MeSH Terms] OR "Mice, SCID"[MeSH Terms] OR "Mice, Congenic"[MeSH Terms] OR "Mice, Inbred Strains"[MeSH Terms] OR "Mice, Mutant Strains"[MeSH Terms] OR "Mice, Obese"[MeSH Terms] OR "Mice, 129 Strain"[MeSH Terms] OR "Mice, Hairless"[MeSH Terms] OR "Rats"[MeSH Terms] OR "Rats, Sprague-Dawley"[MeSH Terms] OR "Rats, Wistar"[MeSH Terms] OR "Rats, Transgenic"[MeSH Terms] OR "Rats, Mutant Strains"[MeSH Terms] OR "Rats, Long-Evans"[MeSH Terms] OR "Rats, Inbred Strains"[MeSH Terms]) NOT ("Adult Germline Stem Cells"[MeSH Terms] OR "Adult Stem Cells"[MeSH Terms] OR "Blastomeres"[MeSH Terms] OR "Cellular Reprogramming Techniques"[MeSH Terms] OR "Embryoid Bodies"[MeSH Terms] OR "Embryonal Carcinoma Stem Cells"[MeSH Terms] OR "Embryonic Germ Cells"[MeSH Terms] OR "Embryonic Stem Cells"[MeSH Terms] OR "Fetal Stem Cells"[MeSH Terms] OR "Granulocyte-Macrophage Progenitor Cells"[MeSH Terms] OR "Hematopoietic Stem Cells"[MeSH Terms] OR "Human Embryonic Stem Cells"[MeSH Terms] OR "Induced Pluripotent Stem Cells"[MeSH Terms] OR "Lymphoid Progenitor Cells"[MeSH Terms] OR "Megakaryocyte-Erythroid Progenitor Cells"[MeSH Terms] OR "Mesenchymal Stem Cells"[MeSH Terms] OR "Mouse Embryonic Stem Cells"[MeSH Terms] OR "Multipotent Stem Cells"[MeSH Terms] OR "Myeloid Progenitor Cells"[MeSH Terms] OR "Myoblasts, Cardiac"[MeSH Terms] OR "Myoblasts, Skeletal"[MeSH Terms] OR "Myoblasts, Smooth Muscle"[MeSH Terms] OR "Myoblasts"[MeSH Terms] OR "Neoplastic Stem Cells"[MeSH Terms] OR "Neural Stem Cells"[MeSH Terms] OR "Oligodendrocyte Precursor Cells"[MeSH Terms] OR "Oogonial Stem Cells"[MeSH Terms] OR "Peripheral Blood Stem Cells"[MeSH Terms] OR "Pluripotent Stem Cells"[MeSH Terms] OR "Precursor Cells, B-Lymphoid"[MeSH Terms] OR "Precursor Cells, T-Lymphoid"[MeSH Terms] OR "Primary Cell Culture"[MeSH Terms] OR "Satellite Cells, Skeletal Muscle"[MeSH Terms] OR "Side-Population Cells"[MeSH Terms] OR "Stem Cell Research"[MeSH Terms] OR "Stem Cells"[MeSH Terms] OR "Thymocytes"[MeSH Terms] OR "Totipotent Stem Cells"[MeSH Terms]) NOT ("3T3

Cells"[MeSH Terms] OR "A549 Cells"[MeSH Terms] OR "BALB 3T3 Cells"[MeSH Terms] OR "Caco-2 Cells"[MeSH Terms] OR "Cell Line, Transformed"[MeSH Terms] OR "Cell Line, Tumor"[MeSH Terms] OR "Cell Line"[All Fields] OR "CHO Cells"[MeSH Terms] OR "COS Cells"[MeSH Terms] OR "HCT116 Cells"[MeSH Terms] OR "HEK293 Cells"[MeSH Terms] OR "HeLa Cells"[MeSH Terms] OR "Hep G2 Cells"[MeSH Terms] OR "HL-60 Cells"[MeSH Terms] OR "HT29 Cells"[MeSH Terms] OR "Jurkat Cells"[MeSH Terms] OR "K562 Cells"[MeSH Terms] OR "L Cells (Cell Line)"[MeSH Terms] OR "LLC-PK1 Cells"[MeSH Terms] OR "Madin Darby Canine Kidney Cells"[MeSH Terms] OR "MCF-7 Cells"[MeSH Terms] OR "NIH 3T3 Cells"[MeSH Terms] OR "PC12 Cells"[MeSH Terms] OR "PC-3 Cells"[MeSH Terms] OR "RAW 264.7 Cells"[MeSH Terms] OR "Sf9 Cells"[MeSH Terms] OR "Swiss 3T3 Cells"[MeSH Terms] OR "THP-1 Cells"[MeSH Terms] OR "Tumor Cells, Cultured"[MeSH Terms] OR "Cells, Cultured"[MeSH Terms] OR "U937 Cells"[MeSH Terms] OR "Vero Cells"[MeSH Terms]) NOT "Computer Simulation"[MeSH Terms] NOT "Humans"[MeSH Terms:noexp] AND (hasabstract[text] AND "animals"[MeSH Terms:noexp] AND English[lang])

"Primary and Stem Cells" (("Adult Germline Stem Cells"[MeSH Terms] OR "Adult Stem Cells"[MeSH Terms] OR "Blastomeres"[MeSH Terms] OR "Cellular Reprogramming Techniques"[MeSH Terms] OR "Embryoid Bodies"[MeSH Terms] OR "Embryonal Carcinoma Stem Cells"[MeSH Terms] OR "Embryonal Carcinoma Stem Cells"[MeSH Terms] OR "Embryonic Germ Cells"[MeSH Terms] OR "Embryonic Stem Cells"[MeSH Terms] OR "Fetal Stem Cells"[MeSH Terms] OR "Granulocyte-Macrophage Progenitor Cells"[MeSH Terms] OR "Hematopoietic Stem Cells"[MeSH Terms] OR "Human Embryonic Stem Cells"[MeSH Terms] OR "Induced Pluripotent Stem Cells"[MeSH Terms] OR "Lymphoid Progenitor Cells"[MeSH Terms] OR "Megakaryocyte-Erythroid Progenitor Cells"[MeSH Terms] OR "Mesenchymal Stem Cells"[MeSH Terms] OR "Mouse Embryonic Stem Cells"[MeSH Terms] OR "Multipotent Stem Cells"[MeSH Terms] OR "Myeloid Progenitor Cells"[MeSH Terms] OR "Myoblasts, Cardiac"[MeSH Terms] OR "Myoblasts, Skeletal"[MeSH Terms] OR "Myoblasts, Smooth Muscle"[MeSH Terms] OR "Myoblasts"[MeSH Terms] OR "Neoplastic Stem Cells"[MeSH Terms] OR "Neural Stem Cells"[MeSH Terms] OR "Oligodendrocyte Precursor Cells"[MeSH Terms] OR "Oogonial Stem Cells"[MeSH Terms] OR "Peripheral Blood Stem Cells"[MeSH Terms] OR "Pluripotent Stem Cells"[MeSH Terms] OR "Precursor Cells, B-Lymphoid"[MeSH Terms] OR "Precursor Cells, T-Lymphoid"[MeSH Terms] OR "Primary Cell Culture"[MeSH Terms] OR "Satellite Cells, Skeletal Muscle"[MeSH Terms] OR "Side-Population Cells"[MeSH Terms] OR "Stem Cell Research"[MeSH Terms] OR "Stem Cells"[MeSH Terms] OR "Thymocytes"[MeSH Terms] OR "Totipotent Stem Cells"[MeSH Terms]) AND "Diseases Category"[MeSH Terms]) NOT ("Clinical Study"[Publication Type] OR "Clinical Trial"[Publication Type] OR "Clinical Trial, Phase I"[Publication Type] OR "Clinical Trial, Phase II"[Publication Type] OR "Clinical Trial, Phase III"[Publication Type] OR "Clinical Trial, Phase IV"[Publication Type])

OR "Clinical Trial, Veterinary"[Publication Type] OR "Clinical Trial Protocol"[Publication Type] OR "Controlled Clinical Trial"[Publication Type] OR "Clinical Conference"[Publication Type] OR "Review"[Publication Type]) NOT ("Animals, Newborn"[MeSH Terms] OR "Disease Models, Animal"[MeSH Terms] OR "Animals, Genetically Modified"[MeSH Terms] OR "Animal Experimentation"[MeSH Terms] OR "Vivisection"[MeSH Terms] OR "Models, Animal"[MeSH Terms] OR "Xenograft Model Antitumor Assays"[MeSH Terms] OR "Neoplasm Transplantation"[MeSH Terms] OR "Leukemia, Experimental"[MeSH Terms] OR "Liver Neoplasms, Experimental"[MeSH Terms] OR "Mammary Neoplasms, Experimental"[MeSH Terms] OR "Melanoma, Experimental"[MeSH Terms] OR "Sarcoma, Experimental"[MeSH Terms] OR "Encephalomyelitis, Autoimmune, Experimental"[MeSH Terms] OR "Myasthenia Gravis, Autoimmune, Experimental"[MeSH Terms] OR "Neuritis, Autoimmune, Experimental"[MeSH Terms] OR "Arthritis, Experimental"[MeSH Terms] OR "Diabetes Mellitus, Experimental"[MeSH Terms] OR "Liver Cirrhosis, Experimental"[MeSH Terms] OR "Neoplasms, Experimental"[MeSH Terms] OR "Nervous System Autoimmune Disease, Experimental"[MeSH Terms] OR "Radiation Injuries, Experimental"[MeSH Terms] OR "Mice"[MeSH Terms] OR "Mice, Knockout"[MeSH Terms] OR "Mice, Transgenic"[MeSH Terms] OR "Mice, Inbred C57BL"[MeSH Terms] OR "Mice, Nude"[MeSH Terms] OR "Mice, SCID"[MeSH Terms] OR "Mice, Congenic"[MeSH Terms] OR "Mice, Inbred Strains"[MeSH Terms] OR "Mice, Mutant Strains"[MeSH Terms] OR "Mice, Obese"[MeSH Terms] OR "Mice, 129 Strain"[MeSH Terms] OR "Mice, Hairless"[MeSH Terms] OR "Rats"[MeSH Terms] OR "Rats, Sprague-Dawley"[MeSH Terms] OR "Rats, Wistar"[MeSH Terms] OR "Rats, Transgenic"[MeSH Terms] OR "Rats, Mutant Strains"[MeSH Terms] OR "Rats, Long-Evans"[MeSH Terms] OR "Rats, Inbred Strains"[MeSH Terms]) NOT ("Organ Culture Techniques"[MeSH Terms] OR "Tissue Culture Techniques"[MeSH Terms] OR "Tissue Extracts"[MeSH Terms] OR "Actihaemyl"[MeSH Terms] OR "Cell Extracts"[MeSH Terms] OR "Liver Extracts"[MeSH Terms] OR "Pancreatic Extracts"[MeSH Terms] OR "Pancreatin"[MeSH Terms] OR "Pancrelipase"[MeSH Terms] OR "Placental Extracts"[MeSH Terms] OR "Thymus Extracts"[MeSH Terms] OR "Isolated Heart Preparation"[MeSH Terms]) NOT ("3T3 Cells"[MeSH Terms] OR "A549 Cells"[MeSH Terms] OR "BALB 3T3 Cells"[MeSH Terms] OR "Caco-2 Cells"[MeSH Terms] OR "Cell Line, Transformed"[MeSH Terms] OR "Cell Line, Tumor"[MeSH Terms] OR "Cell Line"[All Fields] OR "CHO Cells"[MeSH Terms] OR "COS Cells"[MeSH Terms] OR "HCT116 Cells"[MeSH Terms] OR "HEK293 Cells"[MeSH Terms] OR "HeLa Cells"[MeSH Terms] OR "Hep G2 Cells"[MeSH Terms] OR "HL-60 Cells"[MeSH Terms] OR "HT29 Cells"[MeSH Terms] OR "Jurkat Cells"[MeSH Terms] OR "K562 Cells"[MeSH Terms] OR "L Cells (Cell Line)"[MeSH Terms] OR "LLC-PK1 Cells"[MeSH Terms] OR "Madin Darby Canine Kidney Cells"[MeSH Terms] OR "MCF-7 Cells"[MeSH Terms] OR "NIH 3T3 Cells"[MeSH Terms] OR "PC12 Cells"[MeSH Terms] OR "PC-3 Cells"[MeSH Terms] OR "RAW 264.7 Cells"[MeSH Terms] OR "Sf9 Cells"[MeSH Terms] OR "Swiss 3T3 Cells"[MeSH Terms] OR "THP-1 Cells"[MeSH Terms] OR "Tumor Cells, Cultured"[MeSH Terms] OR "Cells, Cultured"[MeSH Terms] OR "U937 Cells"[MeSH Terms] OR "Vero Cells"[MeSH

Terms]) NOT "Computer Simulation"[MeSH Terms] NOT "Humans"[MeSH Terms:noexp] AND (hasabstract[text] AND "animals"[MeSH Terms:noexp] AND English[lang])

"Immortal/Tumor Cell Lines" (("3T3 Cells"[MeSH Terms] OR "A549 Cells"[MeSH Terms] OR "BALB 3T3 Cells"[MeSH Terms] OR "Caco-2 Cells"[MeSH Terms] OR "Cell Line, Transformed"[MeSH Terms] OR "Cell Line, Tumor"[MeSH Terms] OR "Cell Line"[All Fields] OR "CHO Cells"[MeSH Terms] OR "COS Cells"[MeSH Terms] OR "HCT116 Cells"[MeSH Terms] OR "HEK293 Cells"[MeSH Terms] OR "HeLa Cells"[MeSH Terms] OR "Hep G2 Cells"[MeSH Terms] OR "HL-60 Cells"[MeSH Terms] OR "HT29 Cells"[MeSH Terms] OR "Jurkat Cells"[MeSH Terms] OR "K562 Cells"[MeSH Terms] OR "L Cells (Cell Line)"[MeSH Terms] OR "LLC-PK1 Cells"[MeSH Terms] OR "Madin Darby Canine Kidney Cells"[MeSH Terms] OR "MCF-7 Cells"[MeSH Terms] OR "NIH 3T3 Cells"[MeSH Terms] OR "PC12 Cells"[MeSH Terms] OR "PC-3 Cells"[MeSH Terms] OR "RAW 264.7 Cells"[MeSH Terms] OR "Sf9 Cells"[MeSH Terms] OR "Swiss 3T3 Cells"[MeSH Terms] OR "THP-1 Cells"[MeSH Terms] OR "Tumor Cells, Cultured"[MeSH Terms] OR "Cells, Cultured"[MeSH Terms] OR "U937 Cells"[MeSH Terms] OR "Vero Cells"[MeSH Terms]) AND "Diseases Category"[MeSH Terms]) NOT ("Clinical Study"[Publication Type] OR "Clinical Trial"[Publication Type] OR "Clinical Trial, Phase I"[Publication Type] OR "Clinical Trial, Phase II"[Publication Type] OR "Clinical Trial, Phase III"[Publication Type] OR "Clinical Trial, Phase IV"[Publication Type] OR "Clinical Trial, Veterinary"[Publication Type] OR "Clinical Trial Protocol"[Publication Type] OR "Controlled Clinical Trial"[Publication Type] OR "Clinical Conference"[Publication Type] OR "Review"[Publication Type]) NOT ("Animals, Newborn"[MeSH Terms] OR "Disease Models, Animal"[MeSH Terms] OR "Animals, Genetically Modified"[MeSH Terms] OR "Animal Experimentation"[MeSH Terms] OR "Vivisection"[MeSH Terms] OR "Models, Animal"[MeSH Terms] OR "Xenograft Model Antitumor Assays"[MeSH Terms] OR "Neoplasm Transplantation"[MeSH Terms] OR "Leukemia, Experimental"[MeSH Terms] OR "Liver Neoplasms, Experimental"[MeSH Terms] OR "Mammary Neoplasms, Experimental"[MeSH Terms] OR "Melanoma, Experimental"[MeSH Terms] OR "Sarcoma, Experimental"[MeSH Terms] OR "Encephalomyelitis, Autoimmune, Experimental"[MeSH Terms] OR "Myasthenia Gravis, Autoimmune, Experimental"[MeSH Terms] OR "Neuritis, Autoimmune, Experimental"[MeSH Terms] OR "Arthritis, Experimental"[MeSH Terms] OR "Diabetes Mellitus, Experimental"[MeSH Terms] OR "Liver Cirrhosis, Experimental"[MeSH Terms] OR "Neoplasms, Experimental"[MeSH Terms] OR "Nervous System Autoimmune Disease, Experimental"[MeSH Terms] OR "Radiation Injuries, Experimental"[MeSH Terms] OR "Mice"[MeSH Terms] OR "Mice, Knockout"[MeSH Terms] OR "Mice, Transgenic"[MeSH Terms] OR "Mice, Inbred C57BL"[MeSH Terms] OR "Mice, Nude"[MeSH Terms] OR "Mice, SCID"[MeSH Terms] OR "Mice, Congenic"[MeSH Terms] OR "Mice, Inbred Strains"[MeSH Terms] OR "Mice, Mutant Strains"[MeSH Terms] OR "Mice, Obese"[MeSH Terms] OR "Mice, 129 Strain"[MeSH Terms] OR "Mice, Hairless"[MeSH Terms]

OR "Rats"[MeSH Terms] OR "Rats, Sprague-Dawley"[MeSH Terms] OR "Rats, Wistar"[MeSH Terms] OR "Rats, Transgenic"[MeSH Terms] OR "Rats, Mutant Strains"[MeSH Terms] OR "Rats, Long-Evans"[MeSH Terms] OR "Rats, Inbred Strains"[MeSH Terms]) NOT ("Organ Culture Techniques"[MeSH Terms] OR "Tissue Culture Techniques"[MeSH Terms] OR "Tissue Extracts"[MeSH Terms] OR "Actihaemyl"[MeSH Terms] OR "Cell Extracts"[MeSH Terms] OR "Liver Extracts"[MeSH Terms] OR "Pancreatic Extracts"[MeSH Terms] OR "Pancreatin"[MeSH Terms] OR "Pancrelipase"[MeSH Terms] OR "Placental Extracts"[MeSH Terms] OR "Thymus Extracts"[MeSH Terms] OR "Isolated Heart Preparation"[MeSH Terms]) NOT ("Adult Germline Stem Cells"[MeSH Terms] OR "Adult Stem Cells"[MeSH Terms] OR "Blastomeres"[MeSH Terms] OR "Cellular Reprogramming Techniques"[MeSH Terms] OR "Embryoid Bodies"[MeSH Terms] OR "Embryonal Carcinoma Stem Cells"[MeSH Terms] OR "Embryonal Carcinoma Stem Cells"[MeSH Terms] OR "Embryonic Germ Cells"[MeSH Terms] OR "Embryonic Stem Cells"[MeSH Terms] OR "Fetal Stem Cells"[MeSH Terms] OR "Granulocyte-Macrophage Progenitor Cells"[MeSH Terms] OR "Hematopoietic Stem Cells"[MeSH Terms] OR "Human Embryonic Stem Cells"[MeSH Terms] OR "Induced Pluripotent Stem Cells"[MeSH Terms] OR "Lymphoid Progenitor Cells"[MeSH Terms] OR "Megakaryocyte-Erythroid Progenitor Cells"[MeSH Terms] OR "Mesenchymal Stem Cells"[MeSH Terms] OR "Mouse Embryonic Stem Cells"[MeSH Terms] OR "Multipotent Stem Cells"[MeSH Terms] OR "Myeloid Progenitor Cells"[MeSH Terms] OR "Myoblasts, Cardiac"[MeSH Terms] OR "Myoblasts, Skeletal"[MeSH Terms] OR "Myoblasts, Smooth Muscle"[MeSH Terms] OR "Myoblasts"[MeSH Terms] OR "Neoplastic Stem Cells"[MeSH Terms] OR "Neural Stem Cells"[MeSH Terms] OR "Oligodendrocyte Precursor Cells"[MeSH Terms] OR "Oogonial Stem Cells"[MeSH Terms] OR "Peripheral Blood Stem Cells"[MeSH Terms] OR "Pluripotent Stem Cells"[MeSH Terms] OR "Precursor Cells, B-Lymphoid"[MeSH Terms] OR "Precursor Cells, T-Lymphoid"[MeSH Terms] OR "Primary Cell Culture"[MeSH Terms] OR "Satellite Cells, Skeletal Muscle"[MeSH Terms] OR "Side-Population Cells"[MeSH Terms] OR "Stem Cell Research"[MeSH Terms] OR "Stem Cells"[MeSH Terms] OR "Thymocytes"[MeSH Terms] OR "Totipotent Stem Cells"[MeSH Terms]) NOT "Computer Simulation"[MeSH Terms] NOT "Humans"[MeSH Terms:noexp] AND (hasabstract[text] AND "animals"[MeSH Terms:noexp] AND English[lang])

"In Silico" ("Computer Simulation"[MeSH Terms] AND "Diseases Category"[MeSH Terms]) NOT ("Clinical Study"[Publication Type] OR "Clinical Trial"[Publication Type] OR "Clinical Trial, Phase I"[Publication Type] OR "Clinical Trial, Phase II"[Publication Type] OR "Clinical Trial, Phase III"[Publication Type] OR "Clinical Trial, Phase IV"[Publication Type] OR "Clinical Trial, Veterinary"[Publication Type] OR "Clinical Trial Protocol"[Publication Type] OR "Controlled Clinical Trial"[Publication Type] OR "Clinical Conference"[Publication Type] OR "Review"[Publication Type]) NOT ("Animals, Newborn"[MeSH Terms] OR "Disease Models, Animal"[MeSH Terms] OR "Animals, Genetically Modified"[MeSH Terms] OR "Animal Experimentation"[MeSH Terms] OR "Vivisection"[MeSH

Terms] OR "Models, Animal"[MeSH Terms] OR "Xenograft Model Antitumor Assays"[MeSH Terms] OR "Neoplasm Transplantation"[MeSH Terms] OR "Leukemia, Experimental"[MeSH Terms] OR "Liver Neoplasms, Experimental"[MeSH Terms] OR "Mammary Neoplasms, Experimental"[MeSH Terms] OR "Melanoma, Experimental"[MeSH Terms] OR "Sarcoma, Experimental"[MeSH Terms] OR "Encephalomyelitis, Autoimmune, Experimental"[MeSH Terms] OR "Myasthenia Gravis, Autoimmune, Experimental"[MeSH Terms] OR "Neuritis, Autoimmune, Experimental"[MeSH Terms] OR "Arthritis, Experimental"[MeSH Terms] OR "Diabetes Mellitus, Experimental"[MeSH Terms] OR "Liver Cirrhosis, Experimental"[MeSH Terms] OR "Neoplasms, Experimental"[MeSH Terms] OR "Nervous System Autoimmune Disease, Experimental"[MeSH Terms] OR "Radiation Injuries, Experimental"[MeSH Terms] OR "Mice"[MeSH Terms] OR "Mice, Knockout"[MeSH Terms] OR "Mice, Transgenic"[MeSH Terms] OR "Mice, Inbred C57BL"[MeSH Terms] OR "Mice, Nude"[MeSH Terms] OR "Mice, SCID"[MeSH Terms] OR "Mice, Congenic"[MeSH Terms] OR "Mice, Inbred Strains"[MeSH Terms] OR "Mice, Mutant Strains"[MeSH Terms] OR "Mice, Obese"[MeSH Terms] OR "Mice, 129 Strain"[MeSH Terms] OR "Mice, Hairless"[MeSH Terms] OR "Rats"[MeSH Terms] OR "Rats, Sprague-Dawley"[MeSH Terms] OR "Rats, Wistar"[MeSH Terms] OR "Rats, Transgenic"[MeSH Terms] OR "Rats, Mutant Strains"[MeSH Terms] OR "Rats, Long-Evans"[MeSH Terms] OR "Rats, Inbred Strains"[MeSH Terms]) NOT ("Organ Culture Techniques"[MeSH Terms] OR "Tissue Culture Techniques"[MeSH Terms] OR "Tissue Extracts"[MeSH Terms] OR "Actihaemyl"[MeSH Terms] OR "Cell Extracts"[MeSH Terms] OR "Liver Extracts"[MeSH Terms] OR "Pancreatic Extracts"[MeSH Terms] OR "Pancreatin"[MeSH Terms] OR "Pancrelipase"[MeSH Terms] OR "Placental Extracts"[MeSH Terms] OR "Thymus Extracts"[MeSH Terms] OR "Isolated Heart Preparation"[MeSH Terms]) NOT ("Adult Germline Stem Cells"[MeSH Terms] OR "Adult Stem Cells"[MeSH Terms] OR "Blastomeres"[MeSH Terms] OR "Cellular Reprogramming Techniques"[MeSH Terms] OR "Embryoid Bodies"[MeSH Terms] OR "Embryonal Carcinoma Stem Cells"[MeSH Terms] OR "Embryonal Carcinoma Stem Cells"[MeSH Terms] OR "Embryonic Germ Cells"[MeSH Terms] OR "Embryonic Stem Cells"[MeSH Terms] OR "Fetal Stem Cells"[MeSH Terms] OR "Granulocyte-Macrophage Progenitor Cells"[MeSH Terms] OR "Hematopoietic Stem Cells"[MeSH Terms] OR "Human Embryonic Stem Cells"[MeSH Terms] OR "Induced Pluripotent Stem Cells"[MeSH Terms] OR "Lymphoid Progenitor Cells"[MeSH Terms] OR "Megakaryocyte-Erythroid Progenitor Cells"[MeSH Terms] OR "Mesenchymal Stem Cells"[MeSH Terms] OR "Mouse Embryonic Stem Cells"[MeSH Terms] OR "Multipotent Stem Cells"[MeSH Terms] OR "Myeloid Progenitor Cells"[MeSH Terms] OR "Myoblasts, Cardiac"[MeSH Terms] OR "Myoblasts, Skeletal"[MeSH Terms] OR "Myoblasts, Smooth Muscle"[MeSH Terms] OR "Myoblasts"[MeSH Terms] OR "Neoplastic Stem Cells"[MeSH Terms] OR "Neural Stem Cells"[MeSH Terms] OR "Oligodendrocyte Precursor Cells"[MeSH Terms] OR "Oogonial Stem Cells"[MeSH Terms] OR "Peripheral Blood Stem Cells"[MeSH Terms] OR "Pluripotent Stem Cells"[MeSH Terms] OR "Precursor Cells, B-Lymphoid"[MeSH Terms] OR "Precursor Cells, T-Lymphoid"[MeSH Terms] OR "Primary Cell Culture"[MeSH Terms] OR "Satellite Cells, Skeletal

Muscle"[MeSH Terms] OR "Side-Population Cells"[MeSH Terms] OR "Stem Cell Research"[MeSH Terms] OR "Stem Cells"[MeSH Terms] OR "Thymocytes"[MeSH Terms] OR "Totipotent Stem Cells"[MeSH Terms]) NOT ("3T3 Cells"[MeSH Terms] OR "A549 Cells"[MeSH Terms] OR "BALB 3T3 Cells"[MeSH Terms] OR "Caco-2 Cells"[MeSH Terms] OR "Cell Line, Transformed"[MeSH Terms] OR "Cell Line, Tumor"[MeSH Terms] OR "Cell Line"[All Fields] OR "CHO Cells"[MeSH Terms] OR "COS Cells"[MeSH Terms] OR "HCT116 Cells"[MeSH Terms] OR "HEK293 Cells"[MeSH Terms] OR "HeLa Cells"[MeSH Terms] OR "Hep G2 Cells"[MeSH Terms] OR "HL-60 Cells"[MeSH Terms] OR "HT29 Cells"[MeSH Terms] OR "Jurkat Cells"[MeSH Terms] OR "K562 Cells"[MeSH Terms] OR "L Cells (Cell Line)"[MeSH Terms] OR "LLC-PK1 Cells"[MeSH Terms] OR "Madin Darby Canine Kidney Cells"[MeSH Terms] OR "MCF-7 Cells"[MeSH Terms] OR "NIH 3T3 Cells"[MeSH Terms] OR "PC12 Cells"[MeSH Terms] OR "PC-3 Cells"[MeSH Terms] OR "RAW 264.7 Cells"[MeSH Terms] OR "Sf9 Cells"[MeSH Terms] OR "Swiss 3T3 Cells"[MeSH Terms] OR "THP-1 Cells"[MeSH Terms] OR "Tumor Cells, Cultured"[MeSH Terms] OR "Cells, Cultured"[MeSH Terms] OR "U937 Cells"[MeSH Terms] OR "Vero Cells"[MeSH Terms]) NOT "Humans"[MeSH Terms:noexp] AND (hasabstract[text] AND "animals"[MeSH Terms:noexp] AND English[lang])

"Other" ("Humans"[MeSH Terms:noexp] AND "Diseases Category"[MeSH Terms]) NOT ("Clinical Study"[Publication Type] OR "Clinical Trial"[Publication Type] OR "Clinical Trial, Phase I"[Publication Type] OR "Clinical Trial, Phase II"[Publication Type] OR "Clinical Trial, Phase III"[Publication Type] OR "Clinical Trial, Phase IV"[Publication Type] OR "Clinical Trial, Veterinary"[Publication Type] OR "Clinical Trial Protocol"[Publication Type] OR "Controlled Clinical Trial"[Publication Type] OR "Clinical Conference"[Publication Type] OR "Review"[Publication Type]) NOT ("Animals, Newborn"[MeSH Terms] OR "Disease Models, Animal"[MeSH Terms] OR "Animals, Genetically Modified"[MeSH Terms] OR "Animal Experimentation"[MeSH Terms] OR "Vivisection"[MeSH Terms] OR "Models, Animal"[MeSH Terms] OR "Xenograft Model Antitumor Assays"[MeSH Terms] OR "Neoplasm Transplantation"[MeSH Terms] OR "Leukemia, Experimental"[MeSH Terms] OR "Liver Neoplasms, Experimental"[MeSH Terms] OR "Mammary Neoplasms, Experimental"[MeSH Terms] OR "Melanoma, Experimental"[MeSH Terms] OR "Sarcoma, Experimental"[MeSH Terms] OR "Encephalomyelitis, Autoimmune, Experimental"[MeSH Terms] OR "Myasthenia Gravis, Autoimmune, Experimental"[MeSH Terms] OR "Neuritis, Autoimmune, Experimental"[MeSH Terms] OR "Arthritis, Experimental"[MeSH Terms] OR "Diabetes Mellitus, Experimental"[MeSH Terms] OR "Liver Cirrhosis, Experimental"[MeSH Terms] OR "Neoplasms, Experimental"[MeSH Terms] OR "Nervous System Autoimmune Disease, Experimental"[MeSH Terms] OR "Radiation Injuries, Experimental"[MeSH Terms] OR "Mice"[MeSH Terms] OR "Mice, Knockout"[MeSH Terms] OR "Mice, Transgenic"[MeSH Terms] OR "Mice, Inbred C57BL"[MeSH Terms] OR "Mice, Nude"[MeSH Terms] OR "Mice, SCID"[MeSH Terms] OR "Mice, Congenic"[MeSH Terms] OR "Mice, Inbred Strains"[MeSH

Terms] OR "Mice, Mutant Strains"[MeSH Terms] OR "Mice, Obese"[MeSH
 Terms] OR "Mice, 129 Strain"[MeSH Terms] OR "Mice, Hairless"[MeSH Terms]
 OR "Rats"[MeSH Terms] OR "Rats, Sprague-Dawley"[MeSH Terms] OR "Rats,
 Wistar"[MeSH Terms] OR "Rats, Transgenic"[MeSH Terms] OR "Rats, Mutant
 Strains"[MeSH Terms] OR "Rats, Long-Evans"[MeSH Terms] OR "Rats, In-
 bred Strains"[MeSH Terms]) NOT ("Organ Culture Techniques"[MeSH Terms]
 OR "Tissue Culture Techniques"[MeSH Terms] OR "Tissue Extracts"[MeSH
 Terms] OR "Actihaemyl"[MeSH Terms] OR "Cell Extracts"[MeSH Terms] OR
 "Liver Extracts"[MeSH Terms] OR "Pancreatic Extracts"[MeSH Terms] OR
 "Pancreatin"[MeSH Terms] OR "Pancrelipase"[MeSH Terms] OR "Placental
 Extracts"[MeSH Terms] OR "Thymus Extracts"[MeSH Terms] OR "Isolated
 Heart Preparation"[MeSH Terms]) NOT ("Adult Germline Stem Cells"[MeSH
 Terms] OR "Adult Stem Cells"[MeSH Terms] OR "Blastomeres"[MeSH Terms]
 OR "Cellular Reprogramming Techniques"[MeSH Terms] OR "Embryoid Bod-
 ies"[MeSH Terms] OR "Embryonal Carcinoma Stem Cells"[MeSH Terms] OR
 "Embryonal Carcinoma Stem Cells"[MeSH Terms] OR "Embryonic Germ Cells"[MeSH
 Terms] OR "Embryonic Stem Cells"[MeSH Terms] OR "Fetal Stem Cells"[MeSH
 Terms] OR "Granulocyte-Macrophage Progenitor Cells"[MeSH Terms] OR "Hematopoi-
 etic Stem Cells"[MeSH Terms] OR "Human Embryonic Stem Cells"[MeSH Terms]
 OR "Induced Pluripotent Stem Cells"[MeSH Terms] OR "Lymphoid Progenitor
 Cells"[MeSH Terms] OR "Megakaryocyte-Erythroid Progenitor Cells"[MeSH
 Terms] OR "Mesenchymal Stem Cells"[MeSH Terms] OR "Mouse Embryonic
 Stem Cells"[MeSH Terms] OR "Multipotent Stem Cells"[MeSH Terms] OR
 "Myeloid Progenitor Cells"[MeSH Terms] OR "Myoblasts, Cardiac"[MeSH Terms]
 OR "Myoblasts, Skeletal"[MeSH Terms] OR "Myoblasts, Smooth Muscle"[MeSH
 Terms] OR "Myoblasts"[MeSH Terms] OR "Neoplastic Stem Cells"[MeSH Terms]
 OR "Neural Stem Cells"[MeSH Terms] OR "Oligodendrocyte Precursor Cells"[MeSH
 Terms] OR "Oogonial Stem Cells"[MeSH Terms] OR "Peripheral Blood Stem
 Cells"[MeSH Terms] OR "Pluripotent Stem Cells"[MeSH Terms] OR "Precursor
 Cells, B-Lymphoid"[MeSH Terms] OR "Precursor Cells, T-Lymphoid"[MeSH
 Terms] OR "Primary Cell Culture"[MeSH Terms] OR "Satellite Cells, Skeletal
 Muscle"[MeSH Terms] OR "Side-Population Cells"[MeSH Terms] OR "Stem
 Cell Research"[MeSH Terms] OR "Stem Cells"[MeSH Terms] OR "Thymo-
 cytes"[MeSH Terms] OR "Totipotent Stem Cells"[MeSH Terms]) NOT ("3T3
 Cells"[MeSH Terms] OR "A549 Cells"[MeSH Terms] OR "BALB 3T3 Cells"[MeSH
 Terms] OR "Caco-2 Cells"[MeSH Terms] OR "Cell Line, Transformed"[MeSH
 Terms] OR "Cell Line, Tumor"[MeSH Terms] OR "Cell Line"[All Fields] OR
 "CHO Cells"[MeSH Terms] OR "COS Cells"[MeSH Terms] OR "HCT116 Cells"[MeSH
 Terms] OR "HEK293 Cells"[MeSH Terms] OR "HeLa Cells"[MeSH Terms] OR
 "Hep G2 Cells"[MeSH Terms] OR "HL-60 Cells"[MeSH Terms] OR "HT29
 Cells"[MeSH Terms] OR "Jurkat Cells"[MeSH Terms] OR "K562 Cells"[MeSH
 Terms] OR "L Cells (Cell Line)"[MeSH Terms] OR "LLC-PK1 Cells"[MeSH
 Terms] OR "Madin Darby Canine Kidney Cells"[MeSH Terms] OR "MCF-7
 Cells"[MeSH Terms] OR "NIH 3T3 Cells"[MeSH Terms] OR "PC12 Cells"[MeSH
 Terms] OR "PC-3 Cells"[MeSH Terms] OR "RAW 264.7 Cells"[MeSH Terms]
 OR "Sf9 Cells"[MeSH Terms] OR "Swiss 3T3 Cells"[MeSH Terms] OR "THP-

1 Cells"[MeSH Terms] OR "Tumor Cells, Cultured"[MeSH Terms] OR "Cells, Cultured"[MeSH Terms] OR "U937 Cells"[MeSH Terms] OR "Vero Cells"[MeSH Terms]) NOT "Computer Simulation"[MeSH Terms] NOT "Animals"[MeSH Terms:noexp] AND (hasabstract[text] AND English[lang])

“Human” ((“Humans”[MeSH Terms:noexp] AND (“Organ Culture Techniques”[MeSH Terms] OR “Tissue Culture Techniques”[MeSH Terms] OR “Tissue Extracts”[MeSH Terms] OR “Actihaemyl”[MeSH Terms] OR “Cell Extracts”[MeSH Terms] OR “Liver Extracts”[MeSH Terms] OR “Pancreatic Extracts”[MeSH Terms] OR “Pancreatin”[MeSH Terms] OR “Pancrelipase”[MeSH Terms] OR “Placental Extracts”[MeSH Terms] OR “Thymus Extracts”[MeSH Terms] OR “Isolated Heart Preparation”[MeSH Terms] OR “Adult Germline Stem Cells”[MeSH Terms] OR “Adult Stem Cells”[MeSH Terms] OR “Blastomeres”[MeSH Terms] OR “Cellular Reprogramming Techniques”[MeSH Terms] OR “Embryoid Bodies”[MeSH Terms] OR “Embryonal Carcinoma Stem Cells”[MeSH Terms] OR “Embryonal Carcinoma Stem Cells”[MeSH Terms] OR “Embryonic Germ Cells”[MeSH Terms] OR “Embryonic Stem Cells”[MeSH Terms] OR “Fetal Stem Cells”[MeSH Terms] OR “Granulocyte-Macrophage Progenitor Cells”[MeSH Terms] OR “Hematopoietic Stem Cells”[MeSH Terms] OR “Human Embryonic Stem Cells”[MeSH Terms] OR “Induced Pluripotent Stem Cells”[MeSH Terms] OR “Lymphoid Progenitor Cells”[MeSH Terms] OR “Megakaryocyte-Erythroid Progenitor Cells”[MeSH Terms] OR “Mesenchymal Stem Cells”[MeSH Terms] OR “Mouse Embryonic Stem Cells”[MeSH Terms] OR “Multipotent Stem Cells”[MeSH Terms] OR “Myeloid Progenitor Cells”[MeSH Terms] OR “Myoblasts, Cardiac”[MeSH Terms] OR “Myoblasts, Skeletal”[MeSH Terms] OR “Myoblasts, Smooth Muscle”[MeSH Terms] OR “Myoblasts”[MeSH Terms] OR “Neoplastic Stem Cells”[MeSH Terms] OR “Neural Stem Cells”[MeSH Terms] OR “Oligodendrocyte Precursor Cells”[MeSH Terms] OR “Oogonial Stem Cells”[MeSH Terms] OR “Peripheral Blood Stem Cells”[MeSH Terms] OR “Pluripotent Stem Cells”[MeSH Terms] OR “Precursor Cells, B-Lymphoid”[MeSH Terms] OR “Precursor Cells, T-Lymphoid”[MeSH Terms] OR “Primary Cell Culture”[MeSH Terms] OR “Satellite Cells, Skeletal Muscle”[MeSH Terms] OR “Side-Population Cells”[MeSH Terms] OR “Stem Cell Research”[MeSH Terms] OR “Stem Cells”[MeSH Terms] OR “Thymocytes”[MeSH Terms] OR “Totipotent Stem Cells”[MeSH Terms] OR “3T3 Cells”[MeSH Terms] OR “A549 Cells”[MeSH Terms] OR “BALB 3T3 Cells”[MeSH Terms] OR “Caco-2 Cells”[MeSH Terms] OR “Cell Line, Transformed”[MeSH Terms] OR “Cell Line, Tumor”[MeSH Terms] OR “Cell Line”[All Fields] OR “CHO Cells”[MeSH Terms] OR “COS Cells”[MeSH Terms] OR “HCT116 Cells”[MeSH Terms] OR “HEK293 Cells”[MeSH Terms] OR “HeLa Cells”[MeSH Terms] OR “Hep G2 Cells”[MeSH Terms] OR “HL-60 Cells”[MeSH Terms] OR “HT29 Cells”[MeSH Terms] OR “Jurkat Cells”[MeSH Terms] OR “K562 Cells”[MeSH Terms] OR “L Cells (Cell Line)”[MeSH Terms] OR “LLC-PK1 Cells”[MeSH Terms] OR “Madin Darby Canine Kidney Cells”[MeSH Terms] OR “MCF-7 Cells”[MeSH Terms] OR “NIH 3T3 Cells”[MeSH Terms] OR “PC12 Cells”[MeSH Terms] OR “PC-3 Cells”[MeSH Terms] OR “RAW 264.7 Cells”[MeSH Terms]

OR "Sf9 Cells"[MeSH Terms] OR "Swiss 3T3 Cells"[MeSH Terms] OR "THP-1 Cells"[MeSH Terms] OR "Tumor Cells, Cultured"[MeSH Terms] OR "Cells, Cultured"[MeSH Terms] OR "U937 Cells"[MeSH Terms] OR "Vero Cells"[MeSH Terms] OR "Computer Simulation"[MeSH Terms])) AND "Diseases Category"[MeSH Terms]) NOT ("Clinical Study"[Publication Type] OR "Clinical Trial"[Publication Type] OR "Clinical Trial, Phase I"[Publication Type] OR "Clinical Trial, Phase II"[Publication Type] OR "Clinical Trial, Phase III"[Publication Type] OR "Clinical Trial, Phase IV"[Publication Type] OR "Clinical Trial, Veterinary"[Publication Type] OR "Clinical Trial Protocol"[Publication Type] OR "Controlled Clinical Trial"[Publication Type] OR "Clinical Conference"[Publication Type] OR "Review"[Publication Type]) NOT ("Animals, Newborn"[MeSH Terms] OR "Disease Models, Animal"[MeSH Terms] OR "Animals, Genetically Modified"[MeSH Terms] OR "Animal Experimentation"[MeSH Terms] OR "Vivisection"[MeSH Terms] OR "Models, Animal"[MeSH Terms] OR "Xenograft Model Antitumor Assays"[MeSH Terms] OR "Neoplasm Transplantation"[MeSH Terms] OR "Leukemia, Experimental"[MeSH Terms] OR "Liver Neoplasms, Experimental"[MeSH Terms] OR "Mammary Neoplasms, Experimental"[MeSH Terms] OR "Melanoma, Experimental"[MeSH Terms] OR "Sarcoma, Experimental"[MeSH Terms] OR "Encephalomyelitis, Autoimmune, Experimental"[MeSH Terms] OR "Myasthenia Gravis, Autoimmune, Experimental"[MeSH Terms] OR "Neuritis, Autoimmune, Experimental"[MeSH Terms] OR "Arthritis, Experimental"[MeSH Terms] OR "Diabetes Mellitus, Experimental"[MeSH Terms] OR "Liver Cirrhosis, Experimental"[MeSH Terms] OR "Neoplasms, Experimental"[MeSH Terms] OR "Nervous System Autoimmune Disease, Experimental"[MeSH Terms] OR "Radiation Injuries, Experimental"[MeSH Terms] OR "Mice"[MeSH Terms] OR "Mice, Knockout"[MeSH Terms] OR "Mice, Transgenic"[MeSH Terms] OR "Mice, Inbred C57BL"[MeSH Terms] OR "Mice, Nude"[MeSH Terms] OR "Mice, SCID"[MeSH Terms] OR "Mice, Congenic"[MeSH Terms] OR "Mice, Inbred Strains"[MeSH Terms] OR "Mice, Mutant Strains"[MeSH Terms] OR "Mice, Obese"[MeSH Terms] OR "Mice, 129 Strain"[MeSH Terms] OR "Mice, Hairless"[MeSH Terms] OR "Rats"[MeSH Terms] OR "Rats, Sprague-Dawley"[MeSH Terms] OR "Rats, Wistar"[MeSH Terms] OR "Rats, Transgenic"[MeSH Terms] OR "Rats, Mutant Strains"[MeSH Terms] OR "Rats, Long-Evans"[MeSH Terms] OR "Rats, Inbred Strains"[MeSH Terms]) NOT "Animals"[MeSH Terms:noexp] AND (hasabstract[text] AND English[lang])

"Invertebrates" "Drosophila" OR "Drosophila melanogaster" OR "Caenorhabditis" OR "Caenorhabditis elegans" OR "Caenorhabditis/analysis" OR "Caenorhabditis/anatomy and histology" OR "Caenorhabditis/chemistry" OR "Caenorhabditis/classification" OR "Caenorhabditis/cytology" OR "Caenorhabditis/drug effects" OR "Caenorhabditis/embryology" OR "Caenorhabditis/enzymology" OR "Caenorhabditis/etiology" OR "Caenorhabditis/genetics" OR "Caenorhabditis/growth and development" OR "Caenorhabditis/immunology" OR "Caenorhabditis/isolation and purification" OR "Caenorhabditis/metabolism" OR "Caenorhabditis/microbiology" OR "Caenorhabditis/parasitology" OR "Caenorhabditis/pathogenicity"

OR "Caenorhabditis/physiology" OR "Caenorhabditis/radiation effects" OR
 "Caenorhabditis/ultrastructure" OR "Caenorhabditis/virology" OR "Caenorhab-
 ditis elegans/analysis" OR "Caenorhabditis elegans/anatomy and histology"
 OR "Caenorhabditis elegans/chemistry" OR "Caenorhabditis elegans/classification"
 OR "Caenorhabditis elegans/cytology" OR "Caenorhabditis elegans/drug ef-
 fects" OR "Caenorhabditis elegans/embryology" OR "Caenorhabditis elegans/enzymology"
 OR "Caenorhabditis elegans/etiology" OR "Caenorhabditis elegans/genetics"
 OR "Caenorhabditis elegans/growth and development" OR "Caenorhabditis
 elegans/immunology" OR "Caenorhabditis elegans/isolation and purification"
 OR "Caenorhabditis elegans/metabolism" OR "Caenorhabditis elegans/microbiology"
 OR "Caenorhabditis elegans/parasitology" OR "Caenorhabditis elegans/pathogenicity"
 OR "Caenorhabditis elegans/physiology" OR "Caenorhabditis elegans/radiation
 effects" OR "Caenorhabditis elegans/ultrastructure" OR "Caenorhabditis ele-
 gans/virology" OR "Drosophila/analysis" OR "Drosophila/anatomy and histol-
 ogy" OR "Drosophila/chemistry" OR "Drosophila/classification" OR "Drosophila/cytology"
 OR "Drosophila/diagnosis" OR "Drosophila/diagnostic imaging" OR "Drosophila/drug
 effects" OR "Drosophila/embryology" OR "Drosophila/enzymology" OR "Drosophila/epidemiology"
 OR "Drosophila/ethnology" OR "Drosophila/etiology" OR "Drosophila/genetics"
 OR "Drosophila/growth and development" OR "Drosophila/immunology" OR
 "Drosophila/isolation and purification" OR "Drosophila/metabolism" OR "Drosophila/microbiology"
 OR "Drosophila/parasitology" OR "Drosophila/pathogenicity" OR "Drosophila/pharmacology"
 OR "Drosophila/physiology" OR "Drosophila/radiation effects" OR "Drosophila/statistics
 and numerical data" OR "Drosophila/ultrastructure" OR "Drosophila/virology"
 OR "Drosophila melanogaster/analysis" OR "Drosophila melanogaster/anatomy
 and histology" OR "Drosophila melanogaster/chemistry" OR "Drosophila melanogaster/classification"
 OR "Drosophila melanogaster/cytology" OR "Drosophila melanogaster/drug
 effects" OR "Drosophila melanogaster/embryology" OR "Drosophila melanogaster/enzymology"
 OR "Drosophila melanogaster/etiology" OR "Drosophila melanogaster/genetics"
 OR "Drosophila melanogaster/growth and development" OR "Drosophila melanogaster/immunology"
 OR "Drosophila melanogaster/isolation and purification" OR "Drosophila melanogaster/metabolism"
 OR "Drosophila melanogaster/microbiology" OR "Drosophila melanogaster/parasitology"
 OR "Drosophila melanogaster/pathogenicity" OR "Drosophila melanogaster/physiology"
 OR "Drosophila melanogaster/radiation effects" OR "Drosophila melanogaster/ultrastructure"
 OR "Drosophila melanogaster/virology"

2 Annotation guidelines

The GoldHamster II corpus is set up to enable the training of machine learning algorithms for the classification of PubMed abstracts regarding the applied method (e.g. research with primary cells).

The aim of the annotation is to identify and to highlight the type of applied method on the basis of PubMed abstracts. The applied methods shall be categorized according to eight labels listed below:

- In vivo experiments in living vertebrates and cephalopods (**in_vivo**)
- Vertebrate organs and tissues (**organs**)
- Vertebrate primary and stem cells (**primary_cells**)
- Vertebrate immortalized and cancer cell lines (**immortal_cell_line**)
- Experiments with humans or human material (**human**)
- Experiments with invertebrates or invertebrate material (excluding cephalopods) (**invertebrate**)
- In silico (**in_silico**)
- Others (**other**)

The respective type of method is named in parenthesis as configured in the tool that we use for annotation (called TeamTat¹). There are general guidelines for the annotation of the documents:

- The annotation will be carried out by means of highlighting the text with the support of the tool.
- Please highlight only the applied method most likely used in the respective study. Please do not annotate (highlight) methods (models) cited in the text in another context, e.g. as background information in the introduction or discussion sections of the abstract. Some articles also cite species in other context, such as in the example below, which describes an in silico (BLAST) method:

*METHODS: In this study we used the **BLAST reciprocal best-hit methodology** to search for DDR orthologs proteins in *Aedes aegypti*. We also provided a comparison between *Ae. aegypti*, *D. melanogaster* and human DDR network. (PMID 31711518)*

- The annotation shall be carried out based on the title and abstract. Indeed, we noticed that in some articles, the proposed method is only cited in the title but not in the abstract.

¹<https://www.teamtat.org/>

- The identified applied method should be annotated and highlighted according to one of the eight labels.
- It is possible to annotate and highlight two or more methods (e.g. if an in vivo experiment in mice and an experiment with immortalized cell lines is conducted). It is necessary to highlight at least one text span for each label, but it is not necessary to annotate all mentions of the same method.
- However, feel free to annotate all mentions of the same methods, if possible, since this is valuable information for us or for the machine learning. These mentions can provide interesting information regarding the location of the proposed mention in the text or its various forms, e.g., “nematode”, “Caenorhabditis elegans”, “C. elegans” and “worms” (all present in PMID 31713185).
- The length of the highlighted text is not relevant, in case of multi-words. For instance, either “pig coronary arteries in vitro” or “pig coronary arteries” are valid annotations, as long as assigned to the correct label (in this case “organs”).
- To annotate the labels and highlight the methods correctly, you are allowed to use Google or any other search strategy, if, for example, abbreviations or names of cell lines and other methods are unknown.
- If the document does not contain an abstract, it should be ignored, i.e., no annotation at all should be carried out.
- If no label can be allocated, the document should be ignored, i.e., no annotation at all should be carried out. However, the status of the document should be set to “annotated” (cf. “Done” button in Figure 5).

For each label, we provide a short description, explain what the annotator should highlight in the text and show one example from the corpus.

2.1 In vivo experiments on vertebrates, cephalopods

Experiments using certain living animals, i.e. vertebrates like mice, rats, pigs, zebrafish, or cephalopods like octopusses. Those animals are protected by the German Animal Protection Act.

What to annotate? Please highlight the described use of living animals.

*Study was carried out on **rats**, divided into four groups; negative control, positive control, standard drug control, and Myristica fragrans extract treated **rats**. (PMID 31241706)*

2.2 Vertebrate organs and tissues

Experiments using organs or tissues of vertebrate animals that were killed in advance.

What to annotate? Please highlight the described organ- or tissue-based method.

This study describes the effect of variable oxygen supply on relaxing responses induced by alpha-calcitonin gene-related peptide (CGRP) and adrenomedullin (AM) on isolated pig coronary arteries in vitro. (PMID 24084221)

2.3 Vertebrate primary and stem cells

Experiments using primary cells or stem cells that were obtained from vertebrate animals. See <https://web.expasy.org/cgi-bin/cellosaurus/search> “finite” or “stem” for examples.

What to annotate? Please highlight the described primary or stem cell method.

Rabbit tracheal ciliated cell culture was established and CBF was determined using high-speed digital imaging methods. (PMID 22508392)

2.4 Vertebrate immortalized (transformed) and cancer cell lines

Experiments using immortal (transformed) and cancer cell lines (from vertebrate animals) that may be obtained from a provider like ATCC (American Type Culture Collection). See <https://web.expasy.org/cgi-bin/cellosaurus/search> “immortal” for examples.

What to annotate? Please highlight the described immortal cell line method.

This study was conducted to investigate the susceptibility of the porcine ileum epithelial cell line, IPI-2I, to different swine enteric CoVs. (PMID 31176408 and Cellosaurus²)

2.5 Experiments with humans or human materials

Experimental research using living humans or human organs, tissues, stem cells, primary cells, immortalized or cancer cells.

²https://web.expasy.org/cellosaurus/CVCL_3826

What to annotate? Please highlight the described use of humans or human material.

*On this account, our study was directed to scrutinize the influence of this triterpenoid on **human hepatocellular cancer cell model Hep3B**. (PMID 31250646)*

2.6 Experiments with invertebrates or invertebrate materials

Experiments using specific animal species, i.e. invertebrates like *Drosophila*, *Caenorhabditis* (C.) *elegans*. Those animals are NOT protected by the Directive 2010/63/EU.

What to annotate? Please highlight the described (invertebrate) animal species.

***Drosophila** CRYPTOCHROME (dCRY) mediates electrophysiological depolarization and circadian clock resetting in response to blue or ultraviolet (UV) light. [...] To examine these issues in in vivo and in ex vivo whole-brain preparations, we generated **transgenic flies** expressing tryptophan mutant dCRYs in the conserved electron transfer chain and then measured neuronal electrophysiological phototransduction and behavioral responses to light. (PMID 31659046)*

2.7 In silico

Experiments with virtual models.

What to annotate? Please highlight the described in silico model.

*A **computational model** equipped with the main immunological features of the sea bass (*Dicentrarchus labrax* L.) immune system was used to predict more effective vaccination in fish. (PMID 28549079)*

2.8 Other

This includes methods and approaches not covered by any of the above labels, such as human observational studies, in chemico research or use of larvae or early embryos (see attached figure)

What to annotate? Please highlight the described method.

OBJECTIVE: To determine any changes in total hospital revisits within 30 days

of discharge after a hospital stay for medical conditions targeted by the Hospital Readmissions Reduction Program (HRRP). DESIGN: Retrospective cohort study. SETTING: Hospital stays among Medicare patients for heart failure, acute myocardial infarction, or pneumonia between 1 January 2012 and 1 October 2015. (PMID 31405902)

3 PMIDs removed in the second rounds

The following PMIDs were removed from the corpus during the additional rounds of annotations.

- r2.1: 31708395, 31659041, 30638522, 30551685, 31697328, 30658692, 31661072, 28237890, 26108956, 22216973, 27870267
- r2.2: 31405409, 31366381, 31181314, 31212245, 27600312, 31116244, 31357966, 31366540, 31374004, 31366539, 31405426, 31147664, 31405410, 31375236, 31366513, 31366563, 31374005, 31251971, 31330234
- r2.3: 29065140, 30580951, 31029149, 26528898, 29728336, 31170779, 24478394, 31171772, 20197146, 30131013, 29304962, 27139488, 31171792, 27246553, 29730007, 30468255, 30473344, 27833150, 31046828, 28582594, 30744668, 29466325, 30808820

4 Statistics of the corpus

Statistics of the corpus according to the rounds, level of agreement, and merge of the cell lines labels, as listed below:

- rounds: eight combinations of the various rounds, i.e., first round (r1), second round, first phase (r21), second round, second phase (r22), and second round, third phase (r23): r1, r1+r21, r1+r22, r1+r23, r1+r21+r22, r1+r21+r23, r1+r22+r23, and r1+r21+r22+r23.
- agreement: either all annotations from the annotators (All), or only the ones with agreement (Agree) from two annotators;
- cl_merge: either separate labels for “immortal_cell_line” and “primary_cells”, or the two labels merged into “cell_lines”.

Table 1: Statistics of the corpus, training set (All). The values are the average over all 10-fold cross validation splits.

	r1	r1+r21	r1+r22	r1+r23
none	148.5	138.6	148.5	148.5
human	292.5	293.4	259.2	288.9
invertebrate	236.7	224.1	236.7	234.9
in_vivo	434.7	433.8	434.7	417.6
other	361.8	341.1	339.3	345.6
primary_cells	170.1	164.7	169.2	150.3
immortal_cell_line	144.0	142.2	123.3	142.2
organs	234.9	232.2	234.9	192.6
in_silico	194.4	190.8	192.6	193.5
cell_lines	280.8	273.6	259.2	259.2

Table 2: Statistics of the corpus, training set (All). The values are the average over all 10-fold cross validation splits.

	r1+r21+r22	r1+r21+r23	r1+r22+r23	r1+r21+r22+r23
none	138.6	138.6	148.5	138.6
human	260.1	289.8	255.6	256.5
invertebrate	224.1	222.3	234.9	222.3
in_vivo	433.8	416.7	417.6	416.7
other	318.6	324.9	323.1	302.4
primary_cells	163.8	144.9	149.4	144.0
immortal_cell_line	121.5	140.4	121.5	119.7
organs	232.2	189.9	192.6	189.9
in_silico	189.0	189.9	191.7	188.1
cell_lines	252.0	252.0	237.6	230.4

Table 3: Statistics of the corpus, training set (Agree). The values are the average over all 10-fold cross validation splits.

	r1	r1+r21	r1+r22	r1+r23
none	22.5	22.5	22.5	22.5
human	152.1	152.1	184.5	153.9
invertebrate	173.7	188.1	173.7	173.7
in_vivo	311.4	313.2	312.3	330.3
other	117.9	122.4	149.4	125.1
primary_cells	54.0	54.0	54.0	59.4
immortal_cell_line	50.4	50.4	54.9	
organs	117.0	117.0	117.0	128.7
in_silico	115.2	118.8	117.9	122.4
cell_lines	102.6	102.6	107.1	113.4

Table 4: Statistics of the corpus, training set (Agree). The values are the average over all 10-fold cross validation splits.

	r1+r21+r22	r1+r21+r23	r1+r22+r23	r1+r21+r22+r23
none	22.5	22.5	22.5	22.5
human	184.5	153.9	186.3	186.3
invertebrate	188.1	188.1	173.7	188.1
in_vivo	314.1	332.1	331.2	333.0
other	153.9	129.6	156.6	161.1
primary_cells	54.0	59.4	59.4	59.4
immortal_cell_line	54.9	55.8	60.3	60.3
organs	117.0	128.7	128.7	128.7
in_silico	121.5	126.0	125.1	128.7
cell_lines	107.1	113.4	117.9	117.9

Table 5: Statistics of the corpus, development and test sets, for the Agree annotations, the r1+r21+r22+r23 dataset, for both set of cell line labels.

	dev	test
none	1.4	1.1
human	10.1	10.6
invertebrate	10.4	10.5
in_vivo	18.4	18.6
other	8.8	9.1
primary_cells	3.4	3.2
immortal_cell_line	3.4	3.3
organs	7.2	7.1
in_silico	7.2	7.1
cell_lines	6.7	6.4

5 Selection of the hyperparameters

We carried out 40 experiments to decide the best set of hyperparameters, for which we tried the following values:

- batch size: 16 and 32;
- learning rate: 1×10^{-5} , 5×10^{-5} , 1×10^{-4} , 5×10^{-4} ;
- epochs: 10, 20, 30, 40, 50.

All experiments were carried out for the following configuration:

- fold 3 of the cross-validation;
- annotations from rounds 1, 2.1, and 2.2;
- only annotations with agreement;
- the two cell line labels merged into one label.

We present three graphs of the results that we obtained, in which we highlight the distinct values for one of the hyperparameters above. For all figures, the value of the f-score which is depicted is the average one across all labels.

In Figure 1, the results for the highest learning rate (5×10^{-4}) were the lowest ones, while top results were obtained for (1×10^{-4}). Figure 2 shows that there was little impact of the batch size in the results, since both values (16 or 32) obtained low and high scores. Finally, Figure 3 present high and low results for all epochs, but the best result was obtained with 20 epochs.

Based on the the above experiments, the highest F-score of 0,8225 was obtained when considering only the following hyperparameters: batch size of 16, learning rate of 1×10^{-4} , and 20 epochs.

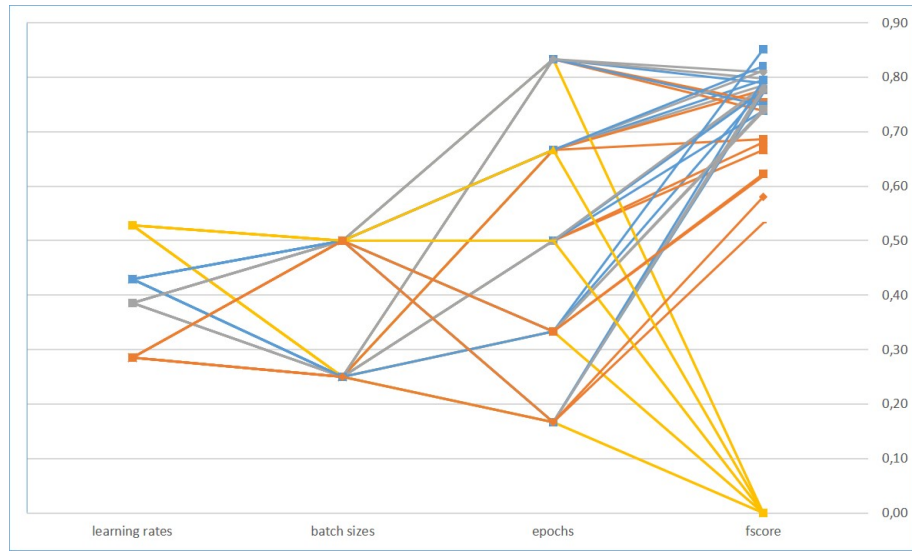


Figure 1: Hyperparameters with focus on the learning rates. The F-score is shown on the right. The values for the three hyperparameters are the following (from bottom to top): (a) learning rates: 1×10^{-5} (bottom), 5×10^{-5} , 1×10^{-4} , 5×10^{-4} (top); (b) batch sizes: 16 (bottom) and 32 (top); and (c) epochs: 10 (bottom), 20, 30, 40, and 50 (top).

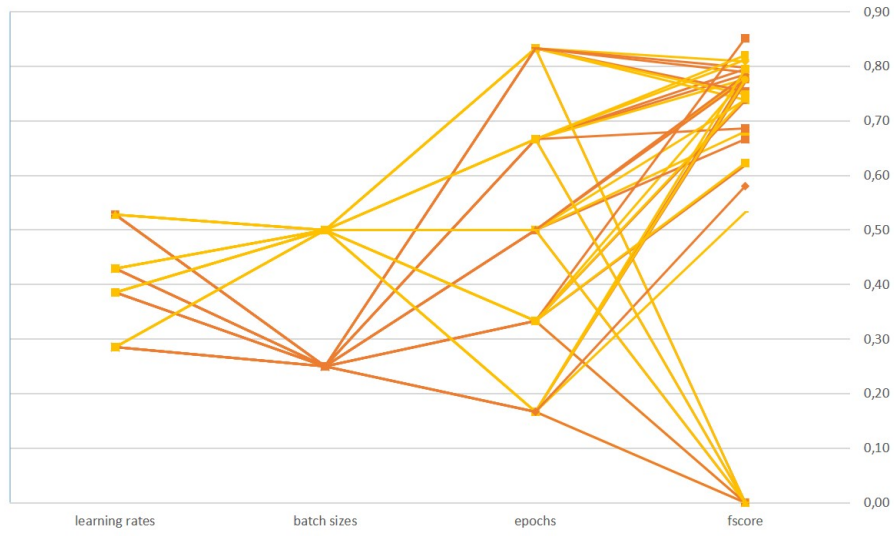


Figure 2: Hyperparameters with focus on the batch sizes. The F-score is shown on the right. The values for the three hyperparameters are the following (from bottom to top): (a) learning rates: 1×10^{-5} (bottom), 5×10^{-5} , 1×10^{-4} , 5×10^{-4} (top); (b) batch sizes: 16 (bottom) and 32 (top); and (c) epochs: 10 (bottom), 20, 30, 40, and 50 (top).

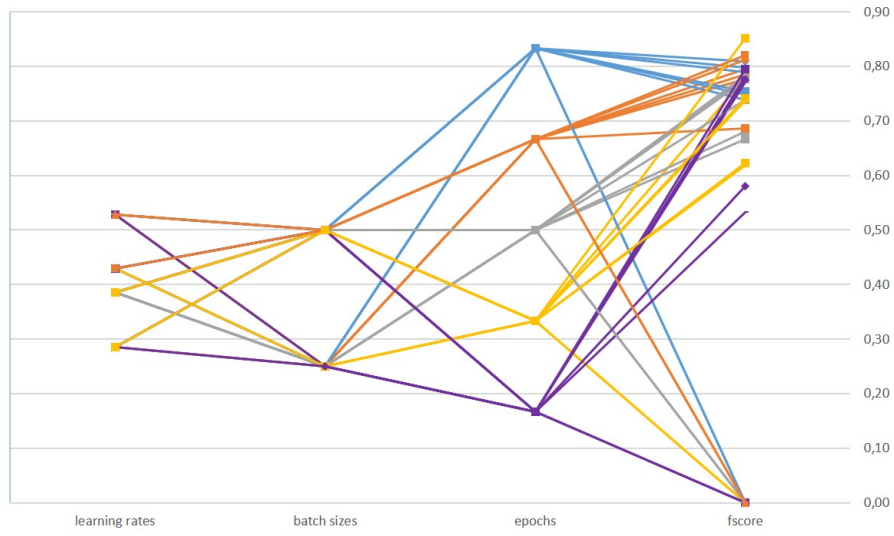


Figure 3: Hyperparameters with focus on the epochs. The F-score is shown on the right. The values for the three hyperparameters are the following (from bottom to top): (a) learning rates: 1×10^{-5} (bottom), 5×10^{-5} , 1×10^{-4} , 5×10^{-4} (top); (b) batch sizes: 16 (bottom) and 32 (top); and (c) epochs: 10 (bottom), 20, 30, 40, and 50 (top).

6 Annotations - error analysis

We list the most frequent labels with disagreements in Tabel 6 below. Of the 20 disagreements listed in this table, 15 of them concern cases where one of the annotators assigned either no label, or the label “others”. This high proportion possibly indicates that annotators might have chosen to assign the “others” label, or no label at all, when encountering an abstract where labeling was unclear. We analyzed each of the disagreements (cf. table) and we summarize below the possible reasons for the misunderstandings:

- Misunderstanding of the annotation guideline, which stipulates to label only according to the actually used experimental model as described in the methodical part of the abstract, but not according to parts of the abstract that constitute related work or background information. [not-used-method]
- Mistake regarding the origin of cell lines. The guideline requires to label cell lines according to source organism: cells from invertebrates as “invertebrates”, while the ones derived from humans as “human” [cell-line-origin]
- Difficulty in identifying whether an animal was killed in advance and its organs or tissues were used subsequently in experiments (label “organs”), or whether the living animal was used for experiments (label “in vivo”). Furthermore, situations in which the authors cite a certain animal organ or tissue in another context, but they did not refer to an ex vivo experiment. [organs-or-invivo]
- Inability to distinguish whether an experiment that involves an animal was indeed an animal experiment according to the EU guidelines. [animal-experim]
- Difficulty in deciding in cases where none of the labels fits an abstract, i.e., whether to assign any label at all, or to assign the label “others” instead. [none-or-others]
- Problem in deciding whether the experimental model described in an abstract corresponds to one of the labels or whether the abstract should be assigned to the “others” label instead. [label-or-others]
- Complexity of the abstract or lack of enough details about the used model. These are cases in which the abstract describes several models and methods without giving enough details, and thus making the identification of a relevant label difficult. [unclear-method]

Table 6: Number of the articles (No.) without full agreement according to their labels and their combinations. We only show the disagreements that occurred at least 10 times. Details about the possible reasons of disagreement were described in details in the text.

No.	ann1	ann2	reason	PMID	disagreement
55	human	others	label-or-others	31357966	case report
46	invivo	invivo, organs	organs-or-invivo	31108095	Hematoxylin-eosin
32	-	others	none-or-others	30151648	review paper
26	human	imm._c.	cell-line-origin	31147664	ovarian cancer cells
22	imm._c.	prim._c.	cell-line-origin	30670030	TMUV-infect. BHK-21
20	invivo	invivo+prim._c.	animal-experim	30556415	embryos
20	human	-	unclear-method	31221068	patient-specific iPSCs
20	human	human+others	label-or-others	31405409	retrospective study
19	others	human+others	label-or-others	31393364	meta-analysis
19	invert.	others	label-or-others	31666534	Drosophila larval
19	invivo	-	unclear-method	22705061	rabbits
18	invivo	others	animal-experim	27246553	dogs
16	insilico	others	label-or-others	29397416	G. hirsutum plants
14	invert.	-	not-used-meth	31661072	Drosophila
13	invivo	organs	organs-or-invivo	30355727	zebrafish heart
12	invivo	invivo+others	none-or-others	31170996	recomb. B. subtilis
12	insilico	-	label-or-others	28237916	parametric study
11	insilico	insilico+others	none-or-others	29032659	A. pleuropneumoniae
10	-	prim._c.	not-used-method	31221064	Mesench. stromal cells
10	invert.	invert.+others	none-or-others	31706086	S. marcescens

7 Evaluation of the additional rounds

We computed the impact of each additional round of annotation in the f-scores for all labels.

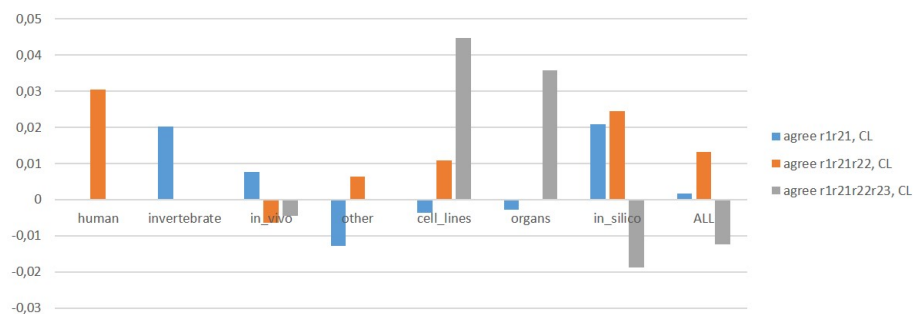


Figure 4: Variation of the f-score for each additional rounds when considering the SVM lineal model. The experiments considered 10-fold cross-validation.

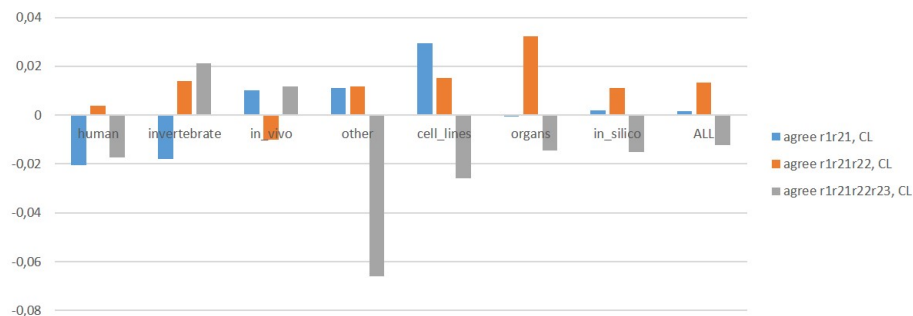


Figure 5: Variation of the f-score for each additional rounds when considering the BioBERT model. The experiments considered 10-fold cross-validation.

8 Analysis of the MeSH terms

We computed the variation on the number of MeSH terms per document according to a range of values for the threshold. This is depicted in the figure below.

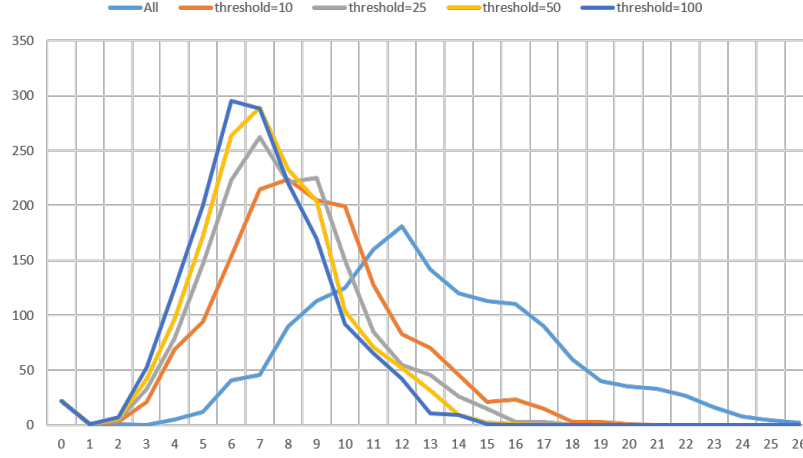


Figure 6: The vertical axis corresponds to the number of documents, and the horizontal axis is the number of Mesh terms per document.

9 Experiments with section labels

We also evaluated whether the performance could improve by concentrating on pre-selected (potentially) relevant parts of the abstracts, e.g., the methods section, instead of the whole abstract. We first considered the sections as provided by PubMed/Medline in structured abstracts. Of our 1,600 articles, only 464 of them were structured.³ We obtained 81 distinct section names and we considered the ones that appeared at least ten times in the abstracts, and which referred to the usual sections in scientific publications. The final list of ten sections that we considered were the following (along with the number of times they occurred): RESULTS (423), CONCLUSIONS (405), BACKGROUND (226), METHODS (364), OBJECTIVE (221), CONCLUSION (22), Results (17), Methods (12), Background (11), and Conclusion (11). This set of sections covers 454 abstracts that contain at least one of the above sections.

With regard to abstracts that were not available in a structured manner, we consider the sections as automatically predicted by the ArguminSci tool⁴. We obtained five labels (along with the number of times they occurred in our collection of 1,600 abstracts), namely: DRI.Background (1565), DRI.Challenge (858), DRI.Approach (1348), DRI.Outcome (1373), DRI.FutureWork (27).

We considered a set of five labels, namely “Background”, “Objective”, “Methods” “Results”, and “Conclusions”, and mapped them to the original ones provided by the structured abstracts or by the ArguminSci tool. Details about the mapping are shown in Table 7.

Table 7: Mapping to the original section names.

Labels	Structured abstracts	ArguminSci
Background	BACKGROUND, Back-ground	DRI.Background
Objective	OBJECTIVE	DRI.Challenge
Methods	METHODS, Methods	DRI.Approach
Results	RESULTS, Results	DRI.Outcome
Conclusions	CONCLUSIONS, CONCLUSION, Conclusion	DRI.FutureWork

We first considered the sections in the structured abstracts, and if not available, we used the ones predicted by ArguminSci. First, we experimented with each of the single section, i.e., the experiments utilized only the text that was included in that particular section. Then, we considered a combination of the sections that obtained the best performance in the first experiments. Similar to the experiments carried out with the MeSH terms (cf. above), we considered the best performing models obtained for the SVM linear and BioBERT. Table 8

³For obtaining the sections from the structured abstracts, we first checked the “NlmCategory” attribute of the abstract text element, and if not found, we considered the “Label” attribute of this element.

⁴<https://github.com/anlausch/ArguminSci>

summarizes the results.

Table 8: Performance (in f-score) when using individual and combined sections for the prediction of the abstract labels. We ran experiments with the SVM linear and the BioBERT models. We show results from considering each section separately, i.e. Background (B), Objective (O), Methods (M), Results (R), and Conclusions (C), as well as combinations of the best performing ones.

	none	B	O	M	R	C	BM	BC	MC	BMC
SVM linear	0.69	0.54	0.51	0.59	0.51	0.60	0.66	0.52	0.59	0.66
BioBERT	0.82	0.73	0.68	0.73	0.65	0.73	0.77	0.79	0.76	0.79

When relying only on some parts of the abstracts, i.e., the sections or their combinations, none of these experiments outperformed the one that relied on the title and whole abstract. However, for both machine-learning methods, some sections achieved a better performance, namely, “Background”, “Methods” and “Conclusions”. For this reason, we only considered combinations of these three best performing sections. In spite of this, we could observe some interesting results for the best combination of sections. For instance, the Background+Methods+Conclusions combination with BioBERT obtained f-score higher than 0.7 for most labels (except for “others”) and over 0.8 for many of them (except for “cell lines” and “others”). Indeed, the “others” label was the ones more affected by the sections, with its f-score dropping from 0.93 (when considering the whole abstract) to 0.62 (for the Background+Methods+Conclusions combination).