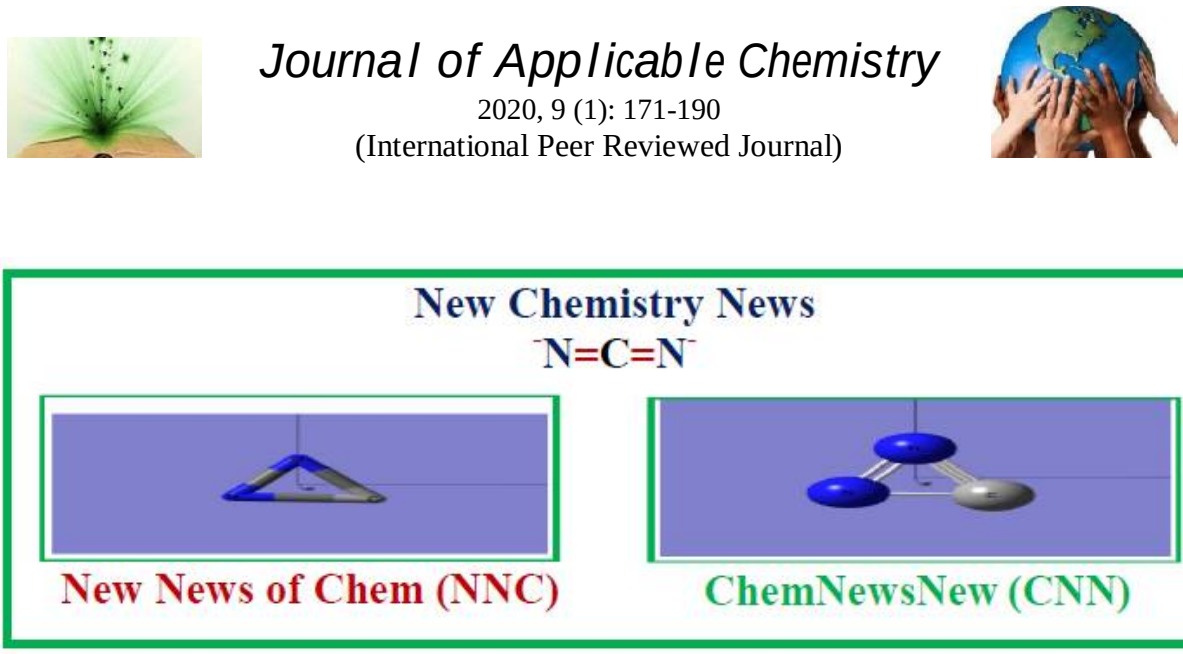
Available online at **www.joac.info**

ISSN: 2278-1862

Journal of Applicable Chemistry

2020, 9 (1): 171-190

(International Peer Reviewed Journal)



Artificial Intelligence (AI)	**Part 1(a) AI. Drug Discovery (Add)** **Part 1(b) AI. Chem.Synth.Org (ACS)**

Task	**AI. applications**
▶ AI-assisted pharmaceutical discovery	▪ Virtual screening o Structure based o Ligand-based ▪ De novo drug design ▪ Property prediction o Physicochemical o Pharmacokinetic ▪ Drug repurposing

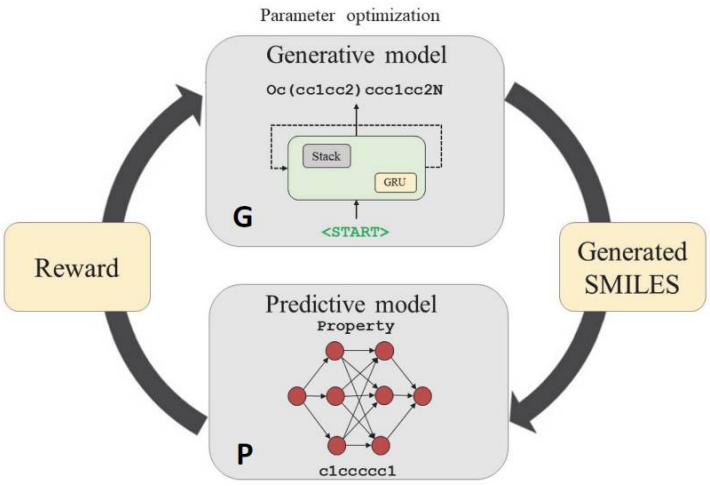
Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery	Chemical Rev, 2019, 119, 18, 10520-10594, DOI: 10.1021/acs.chemrev.8b00728
Xin Yang, Yifei Wang, Ryan Byrne, Gisbert Schneider and Shengyong Yang	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Applications of AI
o Organicsynthesis ▪ HTS in vitro ▪ ADMET-SXR in silico ▪ Iterative cycle to improve o Functional properties of drug candidates

Computer-calculated compounds, Researchers are deploying artificial intelligence to discover drugs	N AT U R E, 557,2018, S 5 5-S57
Nic Fleming	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Methods	o Deep NNs o Generative + Predictive	**Learning**	o ReLeaSE (Reinforcement Learning for Structural Evolution)
Structure input	o SMILES strings	**Model interpretation**	o Recurrent NNs o Adversarial autoencoder
			**Alg.** o First phase: Supervised training of generative and predictive models separately o Second phase: ReLeaSE training of both models jointly o Generation of new chemical structures (feasible SMILES strings) Biased towards desired physical/biological properties o Predictive models o Forecast of desired properties of generated novo–compounds
Fourth industrial revolution	o Combination of big data + AI (Knowledge + Robots) ▪ World economic forum opines a transformation in scientific discovery efforts in future		

Deep reinforcement learning for de novo drug design — Sci. Adv. 2018;4: eaap7885, DOI: 10.1126/sciadv.aap7885

Mariya Popova, OlexandrIsayev, Alexander Tropsha

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Methods	o SVM o Fully connnected NNs o Convolutional NNs o Recurrent NNs	**Learning**	o Deep learning
Software	o Tensorflow o Caffe o Pytorch o Keras o Theano	**Hard-ware**	o GPU o TPU
Feature	**Advantage**	**Feature**	**Advantage**
Rectified linear unit	Avoids vanishing gradients	o Convolutional layers o Pooling layers	Large numbers of input variables
o Dropout o Dropconnect	Surmounts overfitting problem		

The rise of deep learning in drug discovery — Drug Discovery Today, 23, 2018, doi.org/10.1016/j.drudis.2018.01.039

Hongming Chen, Ola Engkvist, Yinhai Wang, Marcus Olivecrona and Thomas Blaschke

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Methods.Chem	o Chemical graph theory o Chemical fingerprints	**Methods.math**	o Deep NNs o Deep Learning o SVM

			o Random forest o Naïve Bayes

Machine learning in chemo informatics and drug discovery	Drug Discovery Today, 2018, doi.org/10.1016/j.drudis.2018.05.010
Yu-Chen Lo, Stefano E. Rensi, Wen Torng and Russ B. Altman	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Goal	o Design of ▪ Automated three-to-five-step synthesis o Prediction 📖 Property 📖 Activity	**Discipline**	Automated chemistry Convergence of artificial intelligence + chemistry → improved drug discovery
Methods	o Recurrent neural networks o SMILES representations (ChEMBL database) o Variational autoencoders o Multitask deep NNs	**Learning**	o One-shot learning o Transfer learning

Evolution of AI assisted automatic synthesis

Level	Design	Synthesis
0	o Manual	o Manual
1	o Input from computational analysis	o Manual
2	o Manual + o Occasional input from AI design	o Manual
3	AI design	Partial automated synthesis + Significant input from human expert
4	AI design	Automated synthesis + occasional input from human expert
5	AI design	Automated synthesis + No (or)minimal input from human expert

The convergence of artificial intelligence and chemistry for improved drug discovery	FutureMed. Chem.,2018.
Clive P Green Ola Engkvist& Garry Pairaudeau	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Methods	Graph neural network	**Learning**	Deep learning
Task	Predicting drug–target interaction	**Training**	DUD-E set 72 proteins
		Testing	25 proteins

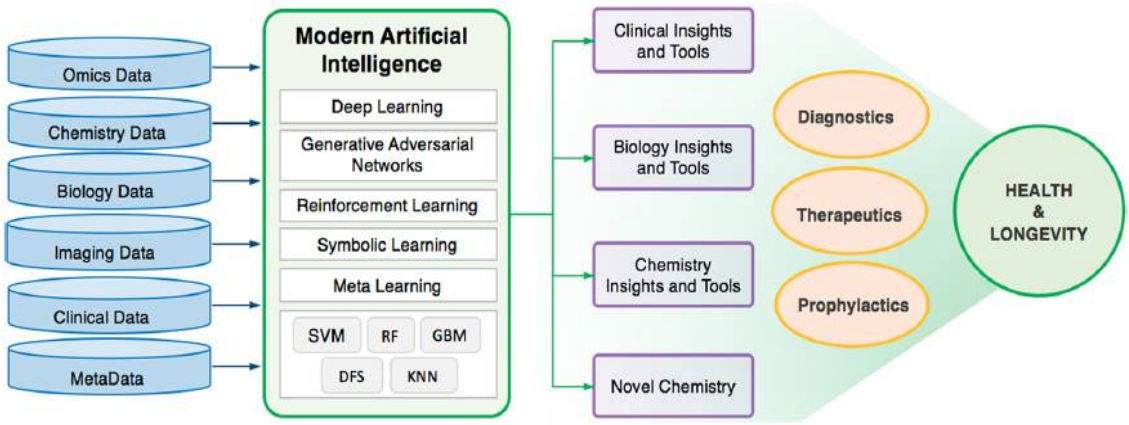
	DUD-E active	DUD-E inactive	PDBbind positive	PDBbind negative
training	15864	973260	1598	9511
test	5841	364149	496	2735

Predicting Drug–Target Interaction Using a Novel Graph Neural Network with 3D Structure-Embedded Graph Representation	J. Chem. Inf. Model. 2019, 59, 3981−3988, DOI: 10.1021/acs.jcim.9b00387

Jaechang Lim, Seongok Ryu, Kyubyong Park, YoJoong Choe, Jiyeon Ham and Woo Youn Kim

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Task	Multidrug resistant tuberculosis patients (230)	Predictio n	**Compound representation**	CNNbased autoencoders >> [voxel-based code; Gaussian blur of atom]
Methods	o Convolutional NN + o Support vector machines o Deep generative models		**DNN performance** .	Image, Voice and Text Recognition, Autonomous Driving, >> Human accuracy



Potential of deep learning for drug discovery

Artificial Intelligence for Drug Discovery, Biomarker Development, and Generation of Novel Chemistry	Mol. Pharmaceutics 2018, 15, 4311−4313, DOI: 10.1021/acs.molpharmaceut.8b00930

Alex Zhavoronkov

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

CQC	o To accelerate development of environmental catalysts	**Applications**	o Captures greenhouse gases (like carbon dioxide)

BASF taps Citrine for catalyst research	CEN.ACS.ORG, 2018

RICK MULLIN

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Compounds	PDB-derived ASTRAL dataset4	**Molecule representation**	Sequence-derived 1D-features
Discipline	AI assisted o Cardiac contractility intervention o Therapeutic disruption of specific protein associations	**Software**	o Python script in Tensor Flow o Run from PyMol 2.3 platform, Schrödinger
Methods	o Deep learning o 2-Hidden Layer NN		

Artificial Intelligence Steering Molecular Therapy in the Absence of Information on Target Structure and Regulation	J. Chem. Inf. Model.,2019, DOI: 10.1021/acs.jcim.9b00651

Ariel Fernández

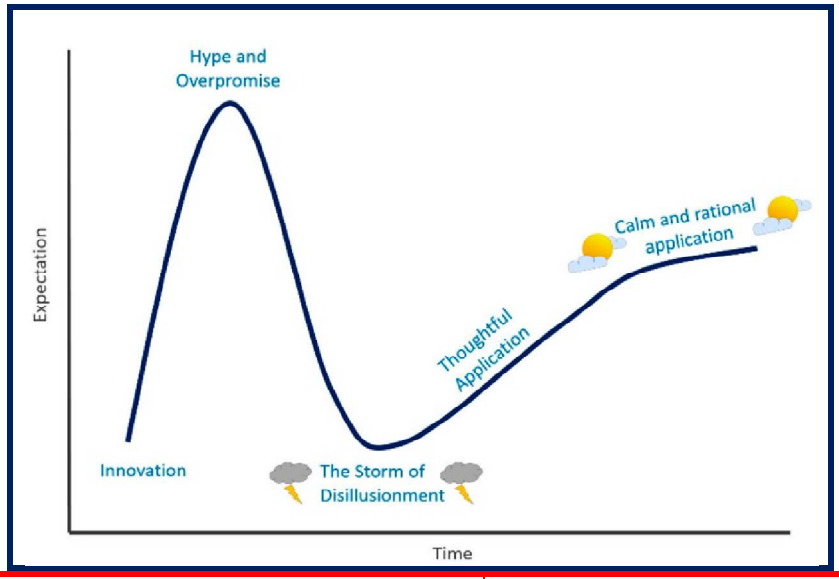
AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Computational models	For prediction + Adjunct + Compliment + Supplement ! Not a panacea	**AI**	🕮 Ultimate goal in drug design ✓ Ability to develop hypotheses ✓ Irrational suggestions based on creativity, insight ✓ Automation, adaptability,

	o If conflicting, to search for correct paradigm o If redundant, no new information		

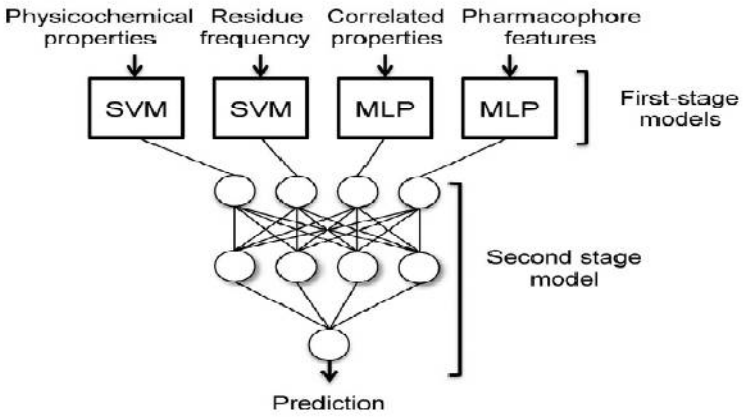
A Novel In Silico Approach to Drug Discovery via Computational Intelligence	J. Chem. Inf. Model. 2009, 49, 1105–1121
David Hecht and Gary B. Fogel	

CI.Drug_Discovery — CI.Discovery_Drug — CI.Discovery — CI.Drug



Artificial Intelligence in Drug Design-The Storm Before the Calm?	ACS Medicinal Chemistry Letters, 2018, 9, 12, 1150-1152, DOI: 10.1021/acsmedchemlett.8b00500
Allan M. Jordan	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Methods	o Deep NN o Restricted Boltzmann NN o Convolutional NN	**Informatics**	o Molecular informatics
Drug Discovery			

Deep Learning in Drug Discovery	Mol. Inf. 2016, 35, 3 – 14, DOI: 10.1002/minf.201501008
Erik Gawehn, Jan A. Hiss, and Gisbert Schneider	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Database	ChEMBL version 24.1	📖 4276 assays 📖 0.5 million compounds

		1.4 million pIC50s	
	NVS data	11 805 assays 1.8 million compounds 18.3 million pIC50s	
Methods	o Profile-quantitative structure−activity relationship (pQSAR)	**Models**	8558 successful models
Testing	▪ pQSAR models updated every month ▪ Predictions	**Applications**	o Virtual screening o Selectivity design o Toxicity promiscuity prediction o Mechanism-of-action rediction
Learning	Transfer learning NVS & ChEMBL		
Alg	Step one: Descriptors: Morgan 2 substructural fingerprints; Models: random forest regression Step two: PLS model		

All-Assay-Max2 pQSAR: Activity Predictions as Accurate as Four-Concentration IC50s for 8558 Novartis Assays	J. Chem. Inf. Model., 2019, DOI: 10.1021/acs.jcim.9b00375
Eric J. Martin, Valery R. Polyakov, Xiang-Wei Zhu, Li Tian, Prasenjit Mukherjee, and Xin Liu	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Task	Computational	Drug discovery
Methods	▪ Machine learning (ML) ▪ Artificial intelligence (AI)	
Data	Big data	
Domain	o Fourth paradigm of science o Fourth industrial revolution ▪ Big data — AI	

Transforming Computational Drug Discovery with Machine Learning and AI	ACS Med. Chem. Let. 2018, 9, 11, 1065-1069, DOI: 10.1021/acsmedchemlett.8b00437
Justin S. Smith, Adrian E. Roitberg, and OlexandrIsayev	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

AI	To design molecules with balance o Potency, selectivity pharmacokinetics	**Pharma company**	GlaxoSmithKline

GSK in computing pact with Exscientia	C&EN Global Enterprise, 2017, 95, 28, 15, DOI: 10.1021/cen-09528-buscon12
LISA JARVIS	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

AI tools	▪ Property prediction ▪ Materials synthesis	**Machine Learning**	o Designing synthesizable molecules with desired range of properties o Pharmaceutical Discovery

Artificial intelligence to spread in drug research	C&EN Global Enterprise 2019, 97, 2, 38-38
Rick Mullin	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Editorial : Special issue on "Artificial Intelligence in DrugDiscovery" in Journal of Medicinal Chemistry	Emphasis: ▪ Impact of AI on drug discovery at present ▪ Not how they might add value in the future ▪ High-quality research yielding "negative" results 🕮 Ex: studies showing no advantages of complex computational methods over simplerapproaches in specific applications.
The Future Is Now: Artificial Intelligence in Drug Discovery	J. Medicinal Chemistry, 2019, 62, 11, 5249, DOI: 10.1021/acs.jmedchem.9b00805
JürgenBajorath, Steven Kearnes, W. Patrick Walters, Gunda I. Georg, Shaomeng Wang,	

AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Task	Prediction	Drug activity		
Compounds	Peptides		**Disease**	Cancer
Database	Chinese traditional medicines (Largest collection)			
Methods	Machine learning models o Random forest o Adaboost regressor o Gradient boosting regressor		**Learning**	Deep Learning models
CQC	300 ns MD simulation			Docking

Artificial Intelligence Approach to Find Lead Compounds for Treating Tumors	J. Phys. Chem. Lett. 2019, 10, 4382−4400 DOI: 10.1021/acs.jpclett.9b01426
Jian-Qiang Chen, Hsin-Yi Chen, Wen-jie Dai, Qiu-JieLv, and Calvin Yu-Chian Chen	

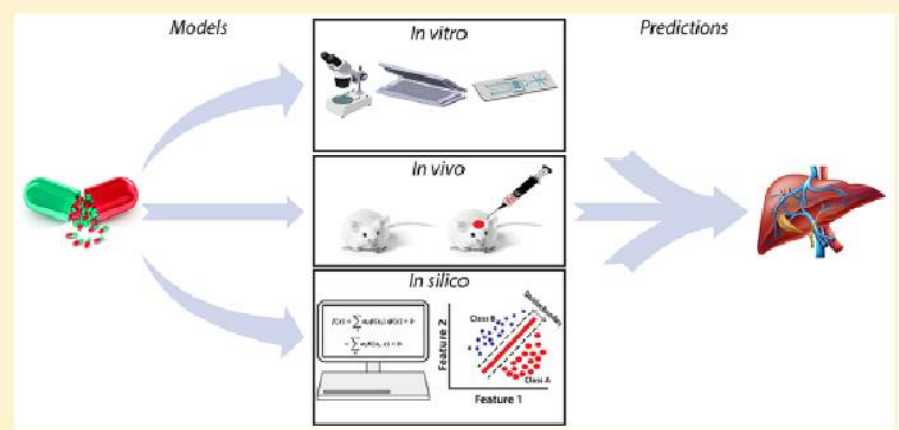
AI.Drug_Discovery — AI.Discovery_Drug — AI.Discovery — AI.Drug

Part 1(b): AI.Chem.Synth.Org

Field	Biomedicine	**Resources & high-end Tools**	▪ IBM's Watson ▪ AI, ML ▪ Cognitive computing resources
Applications	Accurate predictions ▶ Disease diagnosis, drug repurposing ▶ Protein−small-molecule interactions ▶ Biochemical pathways ▶ Human toxicology		
Tools	**Applications**		
AI, ML, cognitive computing	▪ To fill vast knowledge gap of between cellular responses in in vitro and hepatotoxic risk in humans ▪ To process electronic health records (EHRs) ▪ To settle claims data		
DeepNNs	🕮 To elucidate species-specific differences between drug-induced liver injury (DILI) in animal models and reported		

	incidents in humans

Task	🕮 Prediction o Hepatotoxicity	**Discipline**	- Adverse drug reactions (ADRs) - Drug-induced liver injury (DILI)



Mixed-learning method(s) ▪ LDA ▪ Naive Bayes ▪ SVM ▪ ▪ Classification trees ▪ Regression ▪ k-nearest Neighbors ▪ Ensemble of classifiers	**Applications of DNN**	Deep neural nets & deep learning ▪ Cheminformatics ▪ Bioinformatics ▪ Drug discovery
	Undirected graph recursive neural network (UGRNN)	Predict DILI more accurately from physicochemical data

Advancing Predictive Hepatotoxicity at the Intersection of Experimental, in Silico, and Artificial Intelligence Technologies	Chem. Res. Toxicol. 31, 2018, 412−430, DOI: 10.1021/acs.chemrestox.8b00054
Keith Fraser, Dylan M. Bruckner, and Jonathan S. Dordick	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Task	o Prediction of reactions by algorithms o Automated lab equipment	**Database**	▪ US patents ▪ Reaxys reactions database.

Unique features	🕮 Little human help 🕮 Artificial intelligence 🕮 Robot synthesises molecules	**Caution**	Aim o Not to eliminate the chemist o Providing more free time for chemists ▪ For creativity, innovation.
Outcome. planning	For a given molecule o Propose synthetic routes o Reaction conditions o Evaluates best path ▪ Number of steps ▪ Predicted yield.	Outcome. Synthesis	▪ Automated synthesis ▪ Robotic arm ▪ Experiment sets up 🕮 Connecting tubes 🕮 Supplying different reagents 🕮 Flow-chemistry modules 🕮 Reactors 🕮 Membrane-based separators
Compounds synthesized	🕮 Aspirin 91% yield 🕮 (*S*)-warfarin 78% yield 🕮 Five drug-related compounds		

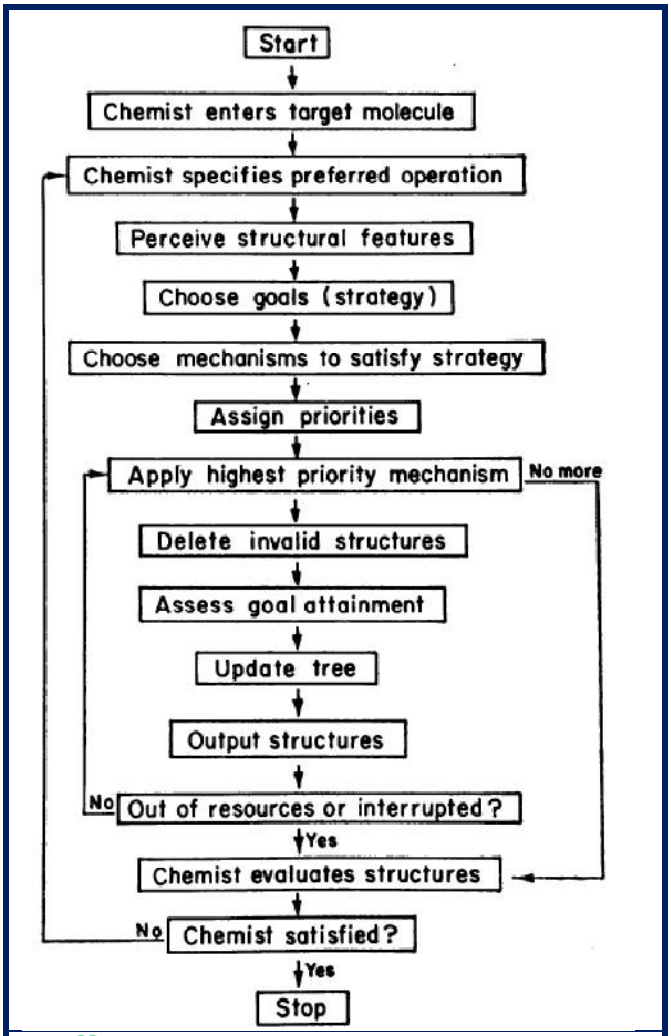
Automating synthesis from planning to execution — CEN.ACS.ORG, 2019.

SAM LEMONICK

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Goal: Organic synthesis

- Tools
  - 🕮 Computer
  - 🕮 Equipment
  - 🕮 Graphical communication



📖 **AI tools**
- 📖 Software
- 📖 Heuristics
  - Chemical knowledge

Computer-Assisted Design of Complex Organic Syntheses	SCIENCE, 166, 1969, 178-192.
E. J. Corey and W. Todd Wipke	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Goal	! Finding better compounds faster 📖 via artificial intelligence	**Collaboration**	✓ Monsanto, drug firms ✓ universities
Sub-goal	50 distinct molecular discovery programs		

AI start-up Atomwise raises $45 million	CEN.ACS.ORG, 2018.
RYAN CROSS	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Organic retro-synthesis (Basis : Logic and Heuristics) o SECS (Simulation and Evaluation of Chemical Synthesis) o SYNCHEM ▪ Improvements: SYNCHEM2 o IGOR (Interactive Generation of Organic Reactions) o KOSP system (Knowledge base-Oriented system for Synthesis Planning)	o STOECH,-- generate automatically all the species by a certain transform. o EROS (Elaboration of Reactions for Organic Synthesis) o SYNGEN + FORWARD o SESAM o CHIRON o LHASA o CAMEO (Computer Assisted Mechanistic Evaluation of Organic) o WODCA (Workbench for the Organisation of Data for Chemical Application)

Computer-aided organic synthesis	Chem. Soc. Rev., 2005, 34, 247–266, DOI: 10.1039/b104620a
Matthew H. Todd	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Task	Organic synthesis with CHEMPUTER	**CHEMPUTER**	🕮 Hardware, software 🕮 Robotic system 🕮 Synthesis abstraction 🕮 Flow chemistry
Outcome	🕮 Leads to a road map ✓ Molecules can be discovered, optimized, made on demand 🕮 Generate new discoveries entirely automatically ✓ Can be verified, optimized, repeated		

Universal Chemical Synthesis and Discovery with 'The Chemputer'	Trends in Chemistry, Month 2019, Doi.org/10.1016/j.trechm.2019.07.004
Piotr S. Gromski, Jarosław M. Granda, and Leroy Cronin	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI

Task	Reaction Prediction	**Training data**	o Polar, hypervalent, radical, pericyclic reactions o Graduate level textbook
Machine Learning	Filtering models o Trained at the level of individual MOs → Reduces space of possible reactions	**Software**	chemoinformatics portalhttp://cdb.ics.uci.edu/
Applications	Retro-synthetic search o Regioselectivity classification o Formability of bonds		

ReactionPredictor: Prediction of Complex Chemical Reactions at the Mechanistic Level Using Machine Learning	J. Chem. Inf. Model., 52, 2012, 2526−2540, dx.doi.org/10.1021/ci3003039
Matthew A. Kayala and Pierre Baldi	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Task	Predict course of arbitrary chemical reactions	**Methods**	o Physical & chemical laws o Rule-based expert systems o Inductive machine learning
Database	Reaction Explorer	**Data**	▪ 1630 full multistep reactions ▪ 2358 distinct starting materials, intermediates ▪ 2989 productive mechanistic steps ▪ 6.14 million unproductive mechanistic steps

Machine learning

1) Atom level reactivity filters trained ones to prune 94.00% of nonproductive reactions with a 0.01% error rate
2) An ensemble of ranking models trained on pairs of interacting MOs → learns a relative productivity function over mechanistic steps of system

CAMEO7	o Complex set of heuristics for different classes ▶ Predicts multistep reactions	o Beppe o Sophia	Multistep reactions –identification of o Reactive sites o Reactions
EROS8	o Uses configurable system composed of multistep reaction o Graph-based rule libraries o Extra modules to add more constraints o Heats of formation, physicochemical properties, kinetic simulations	Reaction Explorer system	o Detailed graph rewrite rules for individual mechanistic steps o Not like a common practice of a single transformation for an overall reaction from starting materials to final products. o Rules described using an alternative physically motivated "electron-flow" Specification → allows visualization of the "arrow-pushing" diagrams for each mechanistic step
ToyChem11 Robia12	o Build on the EROS idea of physicochemical constraints		

Limitatios

- Curation of large amounts of
  - o Rules, exception handling
  - o Expert knowledge
- Unmanageable at larger scales
- Adding/modifications of new existing ancetedents/consequent difficult
- Lack generality
  - **If** a particular reaction pattern not encoded explicitly,
    Then system will never be able to return the corresponding reaction.

Rule-based reaction prediction

- o Knowledge-based
- o Human encoding of heuristics
- o Graph-rewrite

- ▪ Computationally tractable
  - o Patterns, and constraints
- ▪ Quick predictions

Learning to Predict Chemical Reactions	J. Chem. Inf. Model., 512011, , 2209–2222, dx.doi.org/10.1021/ci200207y
Matthew A. Kayala, Chlo e-Agathe Azencott, Jonathan H. Chen, and Pierre Baldi	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Task	o Molecular design o Optimization	**Discipline**	o Cell metabolism o Prebiotic chemistry
Methods	**Chemical heuristics**	→ Path of traversing high-dimensional reactive PES	
	CQC structure optimizations	→ 3D-geometric structures → Eenergies of the products & intermediates	
Test Case study	**Heuristic-CQC computations**reproduce experimentally observed reaction products, major reaction pathways Ex: autocatalytic cycles of formose reaction i.e. self-condensation of formaldehyde in alkaline solution and at surfaces of various minerals		

Complex Chemical Reaction Networks from Heuristics-Aided Quantum Chemistry	J. Chem. Theory Comput., 10, 2014, 897−907, dx.doi.org/10.1021/ct401004r
DmitrijRappoport, Cooper J. Galvin, Dmitry Yu. Zubarev, and AlánAspuru-Guzik	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

LHASA-1 (Logic and Heuristics Applied to Synthetic Analysis)	- lacks stereochemical capability modules

Computer-assisted Analysis of Complex Synthetic Problems (Review)	Q. Rev. Chem. Soc., 1971,25, 455-482
E. J. Corey	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

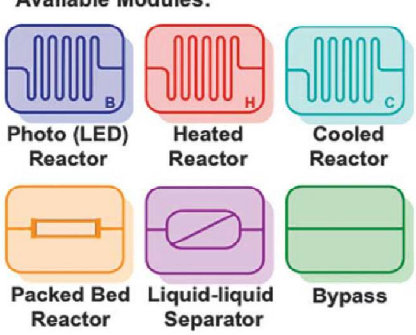
Task	Conformations (off-equilibrium) calculation	**Database**	Large computational DFT database Database17 o Contains 166.4Billion molecules containing up to 17 atoms of C, N, O, S, and halogens
Compounds	20 Million subset: 57,462 small organic molecules	**Future**	Development of future general-purpose machine learned CQC potentials

Data Descriptor: ANI-1, A data set of 20 million calculated off-equilibrium conformations for organic molecules	Sci. Data 4:170193 doi: 10.1038/sdata.2017.193 (2017).
Justin S. Smith, OlexandrIsayev& Adrian E. Roitberg	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI

Auto-optimized organic synthesis system

- **Unique features**
  - Fully integrated
  - Versatile
  - Reconfigurable



Modules

- Heated reactor (up to 120°C)
- Cooled reactor (to –20°C),
- Light-emitting diode (LED)–based photochemistry
- Photochemistry reactor
- Packed-bed reactor (for solid supported reagents; catalysts, passive mixing)
- Membrane-based liquid-liquid separator.(purification via extraction)
- Bypass (for reagent addition in a minimal volume, mixing, unused bay)

Applications

🕮 Automated optimization of a specific reaction or sequence of reactions 🕮 Synthesis of a range of substrates under user-selected conditions, ex: Scope of transformation under optimum conditions 🕮 Scale-up of a selected synthesis from a previous optimum conditions

Reconfigurable system for automated optimization of diverse chemical reactions	Science 361, 2018, 1220–1225.
Anne-Catherine Bédard, Andrea Adamo, Kosi C. Aroh, M. Grace Russell, Aaron A. Bedermann, Jeremy Torosian, Brian Yue, Klavs F. Jensen, Timothy F. Jamison	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

🕮 Forward-reaction prediction 🕮 IBM's innovative approach	Input: starting materials Method: 395 496 reactions trained with NN Output: Prediction of reaction under new experimentalconditions. Figure of merit: Correct 80 percent of the time
Chematica	Computer program capable of designing novel efficient syntheses of medicinally relevant molecules
Manual organic Synthesis o Hand coded reaction scheme o Execution in lab o Time consuming task o Nonoptimized solutions (frequent)	**Synthesis and retrosynthesis** 🕮 Corey and Wipke o Envisioned machine design using handcrafted rules (or reaction templates) - Writing rules remained a time-consuming task o Remedy: Deep chemical expertise

Analogy between

	Organic synthesis	**Solitaire game**
Pieces on the board in Solitaire game	Beginning of game	Precursor molecules
Winning game	Only one piece Remains	Target Molecule

Google's AI program Alphazero	**Chess Human player** + Takes few minutes to learn to play - Decades of lifetime to become a master **Alphazero** Input: Only rules of chess Learning: Few hours of self-learning Outcome: ! Adventurous and unconventional way of playing ! Beats human masters or existing programs
Future of chemical synthesis; drug discovery; medical diagnosis	✓ Chemistry is more complex than chess game ! Machine learning + Deep NN + Heuristics → Speed up drugs development

Artificial Intelligence: The Future for Organic Chemistry?	ACS Omega 3, 2018, 13263−13266, DOI: 10.1021/acsomega.8b01773
Franck Peiretti and Jean Michel Brunel	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI

Task	Retrosynthesis	**Database**	Reaxys36 chemistry database o 12.4 million single-step reactions o Two sets of extracted rules
Methods	Retrosynthetic routes discovery o Monte carlotree search o Symbolic AI	**NNs**	Deep highway network with bonds

Heuristic Methods for chemical synthesis

1970-2010	**Now (This decade)**	**Training: reaction centre**
o Expert chemists dictated hand-coded heuristics (rules) of chemistry to computers	o Strong, general planning algorithms o Symbolic representations o Autonomous learning o Rich history of chemistry ➔ Accepting machine as an assistant in chemical synthesis	o Trainig: reactions published before 2015 o Validation & testing: reactions data from 2015 onwards o Rollout policy network: SLP, trained with 17,134 rules o **Applications** o Agriculture o Healthcare o Material science.

Planning chemical syntheses with deep neural networks and symbolic AI	N AT U R E, 555,2018, 604-618 doi:10.1038/nature25978
Marwin H. S. Segler, Mike Preuss & Mark P. Waller	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Task	To assess protective groups (PGs) reactivity as a function ofreaction conditions (catalyst, solvent)	**Database**	Reaxys database o Catalytic hydrogenation reactions
Methods	🕮 Condensed Graph of Reaction (CGR)	**Data**	▪ Chemical transformations proceeding under ca. 271000 reaction conditions

Automatized assessment of protective group reactivity: a step toward big reaction data analysis	J. Chem. Inf. Model.,2016,DOI: 10.1021/acs.jcim.6b00319
Arkadii I. Lin, Timur IsmailovichMadzhidov, Olga Klimchuk, Ramil I. Nugmanov, Igor S. Antipin, and Alexandre Varnek	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI

Task	AI based organic synthesis	**Database**	Granted US patents o 15 000 experimental reaction records

Output	Major product		

Prediction of Organic Reaction Outcomes Using Machine Learning	ACS Cent. Sci., 2017, DOI: 10.1021/acscentsci.7b00064
Connor W. Coley, Regina Barzilay, Tommi S. Jaakkola, William H. Green, and Klavs F. Jensen	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI

Learning.Mach	
Pharma research o Physicochemical property prediction o Formulation prediction, o ADME/tox o Target prediction o Skin permeation	**Commonman's daily use** o Internet searches o Voice recognition o Vision software o Phones, cameras o Self-driving cars. o Robots. Smartphones o Voice recognition software like SIRI o Read the news o Make a purchase on the internet via AMAZON o Use social network software o Large Companies Baidu, Google, Facebook etc. 🕮 Use deep learning in facial recognition algorithms alone

Learning.Mach. Cheminformatics	o SVM o k-Nearest Neighbors o Naïve Bayesian o Decision Trees	**Tasks**	o Binary classification o Multiple classes

The Next Era: Deep Learning in Pharmaceutical Research	Pharm Res., 2016, DOI 10.1007/s11095-016-2029-7
Sean Ekins	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI

Task	Drug discovery
Traditional experiments - Expensive - time consuming	Modern approach ▪ Machine learning approaches evolved into deep learning ▪ Big data ▪ High computing power
Machine learning o LDA o SVM o DT RF o kNN, o ANN	Deep learning ▪ CNN ▪ DNN ▪ RNN ▪ DBN
Big' data o Volume (scale of data) o Velocity (growth of data) o Variety (diversity of sources) o Veracity (data uncertainty)	Limitations. Deep learning - Availability of a large amount of high-quality data - Non availability of biomedical data generated by pharmaceutical companies to academic institutesor public - Lack of rational interpretations of associated biological mechanisms - Black box mode of 'Deep learning models

Commercial drugs and drug candidates discovered by computational methods

Computational Method	Drug function	Drug	Year
Docking	Inhibits Checkpoint kinase 1	CCT244747	2012
SAR/QSAR	Inhibits hepatitis C RNA replication	PTC725	2014
SAR/QSAR	Treats spinal muscular atrophy	RG7800	2016
Molecular modeling	Inhibits phosphatidylinositol-3-kinase	GDC-0941	2015

From machine learning to deep learning: progress in machine intelligence for rational drug discovery — Drug Discovery Today, 22, 2017, http://dx.doi.org/10.1016/j.drudis.2017.08.010

Lu Zhang, Jianjun Tan, Dan Han and Hao Zhu

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Chematica program	▪ Autonomous design of synthetic pathways	**Application**	o Seven structurally diverse targets + One natural product
Figure of merit	o Computer generated synthetic schemes were successfully executed in chemical laboratory + Offer significant yield + Improvements cost savings + Provide alternatives to patented routes + Produced targets that were not synthesized previously		
Future targets	▪ In silico colleague [multiprocessor machines potentially linked into larger clusters] o Constantly learns o Never forgets, o Will never retire ▪ Syntheses of very complex targets		

Efficient Syntheses of Diverse, Medicinally Relevant Targets Planned by Computer and Executed in the Laboratory — Chem 4, 2018, 522–532, doi.org/10.1016/j.chempr.2018.02.002

Tomasz Klucznik, Barbara Mikulak-Klucznik, Michael P. McCormack, Heather Lima, Sara Szymku, ManishabrataBhowmick, Karol Molga, Yubai Zhou, Lindsey Rickershauser, Ewa P. Gajewska, Alexei Toutchkine, Piotr Dittwald, Michał P. Startek, Gregory J. Kirkovits, Rafał Roszak, Ariel Adamski, BiankaSieredzinska, Milan Mrksich, Sarah L.J. Trice, and Bartosz A. Grzybowski

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Goal	▶ Insilico Medicine	**Task**	▶ Hunting for small molecules [inhibitors of discoidin domain receptor 1 (DDR1)]
Database	▶ DDR1 inhibitors ▶ Kinase inhibitors ▶ Nonkinase inhibitors ▶ Patent-protected molecules	**Generative Reinforcement learning**	▶ Uses rewards (scored for molecules that satisfy its goals) to guide the algorithm

- Time period: 46 days
- The alg. proposed 30000 potential drugs
- Computer software filtered compounds
- Chemists selected six molecules
- Two of them showed no activity

- Two other molecules with promising activity were synthesized in lab
  - One compound performed well against kinase screens
  - Its metabolic stability in mice tested

Analogy: Alg. to find Times Square in New York City using Google Maps

AI identifies drug candidate in weeks	CEN.ACS.ORG, 2019.
SAM LEMONICK	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Task	Predictive models	Reactive trajectories	Prediction of reactivity o Conformational descriptors alone Accuracy:85%
Methods	o Computational Statistical mechanics—transition interface sampling		▪ Simulates Kinetics of reaction
	o LASSO		▪ Feature regularization
	o QM/MM TIS		▪ Generated reactive trajectories
	o Machine learning		▪ Selects features relevant to reactivity

Machine Learning Identifies Chemical Characteristics That Promote Enzyme Catalysis	J. Am. Chem. Soc., 141, 2019, 4108−4118,DOI: 10.1021/jacs.8b13879
Brian M. Bonk, James W. Weis, and Bruce Tidor	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Methods	o Molecularmatched pair (MMP) analysis	**Discipline**	QSAR.special case

Application. Bio

- o ADME
- o Bioisosterism
- o Plasma protein binding
- o Oral exposure
- o Potency
- o Intrinsic clearance
- o Metabolism
  - ▪ Herg and p450 in vitro
- o Glucuronidation clearance
- o Selectivity against off-targets
- o Mode of action

Application.Properties

- ▪ Aqueous solubility
- ▪ logD, lipophilicity

Matched Molecular Pair Analysis in Short: Algorithms, Applications and Limitations	Computational and Structural Biotechnology Journal 15 (2017) 86–90, http://dx.doi.org/10.1016/j.csbj.2016.12.003
Christian Tyrchan, Emma Evertsson	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Task	Hospital mortality	Predicting	**Learning**	Deep learning models
Database	Electronic health records from two US academic medical centers		**Data**	▪ Adult patients: 216,221 ▪ Data unrolled → 46,864,534,945 points ▪ Discharge diagnoses :1–228

Scalable and accurate deep learning with electronic health records	Npj (Nature Partner Journals) Digital Medicine, 2018, 1:18 , doi:10.1038/s41746-018-0029-1
Alvin Rajkomar, Eyal Oren, Kai Chen, Andrew M. Dai, Nissan Hajaj, Michaela Hardt , Peter J. Liu , Xiaobing Liu , Jake Marcus, Mimi Sun , Patrik Sundberg, Hector Yee, Kun Zhang, Yi Zhang, Gerardo Flore , Gavin E. Duggan, Jamie Irvine, Quoc Le, Kurt Litsch, Alexander Mossin , Justin Tansuwan, De Wang, James Wexler, Jimbo Wilson, Dana Ludwig , Samuel L. Volchenboum, Katherine Chou, Michael Pearson, Srinivasan Madabushi, Nigam H. Shah, Atul J. Butte , Michael D. Howell, Claire Cui, Greg S. Corrado and Jeffrey Dean	

AI.Sythesis_Organic —- AI.Sythesis —- AI.Organic —- AI.

Deep architectures and deep learningin chemoinformatics: the prediction of aqueous solubility for druglikemolecules	J Chem Inf Model. 2013;53(7):1563–75
Lusci A, Pollastri G, Baldi P.	

Deep learning for druginducedliver injury	J Chem Inf Model. 2015;55(10):2085–93
Xu Y, Dai Z, Chen F, Gao S, Pei J, Lai L.	

Deep neuralnetsas amethod for quantitative structure-activity relationships	J ChemInf Model. 2015;55(2):263–74
Ma J, Sheridan RP, Liaw A, Dahl GE, Svetnik V.	

Deep biomarkers of human aging: application of deep neural networks to biomarker development	Aging (AlbanyNY) 2016;8(5):1021–33
Putin E, Mamoshina P, Aliper A, Korzinkin M, Moskalev A,Kolosov A	

Deep architectures for protein contactmap prediction	Bioinformatics. 2012;28(19):2449–57
Di Lena P, Nagata K, Baldi P.	

Using deep learning toenhance cancer diagnosis and classification	In: Proceeding of the30th International conference on machine learning. Atlanta, GA: JMLR: W&CP; 2013.
Fakoor R, Ladhak F, Nazi A, Huber M.	

Deep convolutional neuralnetworks for annotating gene expression patterns in the mousebrain	BMC Bioinf. 2015;16:147
Zeng T, Li R, Mukkamala R, Ye J, Ji S.	

A deep learning framework for modeling structural features of RNAbinding protein targets. Nucleic	Acids Res. 2016;44(4):e32
Zhang S, Zhou J, Hu H, Gong H, Chen L, Cheng C	

Modeling epoxidation of drug-like molecules with a deep machine learning network	ACS Cent Sci. 2015;1(4):168–80
Hughes TB, Miller GP, Swamidass SJ.	

DeepTox: toxicity prediction using deep learning.	Front Environ Sci. 2016;3:80
Mayr A, Klambauer G, Unterthiner T, Hochreiter S.	

Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data	MolPharm. 2016;13(7):2524–30.
Aliper A, Plis S, Artemov A, Ulloa A, Mamoshina P, Zhavoronkov A.	

Deep learning in label-free cell classification	Sci Rep. 2016;6:21471
Chen CL, Mahjoubfar A, Tai LC, Blaby IK, Huang A, NiaziKR,et al.	

Classifying and segmenting microscopy images with deep multiple instance learning	Bioinformatics. 2016;32(12):i52–9
Kraus OZ, Ba JL, Frey BJ.	

Intra- and inter-fractional variation prediction of lung tumors using fuzzydeep learning	IEEE JTranslEng Health Med. 2016;4:4300112.
Park S, Lee SJ, Weiss E, Motai Y.	

ACS.org ;sciencedirect.com : Information Source

R. Sambasiva Rao, School of Chemistry
Andhra University, Visakhapatnam
rsr.chem@gmail.com