

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# Computational and Structural Biotechnology Journal

journal homepage: [www.elsevier.com/locate/csbj](http://www.elsevier.com/locate/csbj)

## Research Article

# Design rules applied to silver nanoparticles synthesis: A practical example of machine learning application.

Irini Furxhi<sup>a,b</sup>, Lara Faccani<sup>a,\*</sup>, Ilaria Zanoni<sup>a,\*</sup>, Andrea Briigliadori<sup>a,1</sup>, Maurizio Vespignani<sup>a</sup>, Anna Luisa Costa<sup>a</sup>

<sup>a</sup> CNR-ISSMC (Former ISTEC), National Research Council of Italy-Institute of Science, Technology and Sustainability for Ceramics, Faenza, Italy

<sup>b</sup> Transgero Limited, Limerick, Ireland



## ARTICLE INFO

### Keywords:

Silver nanoparticles, synthesis, machine learning  
Shapley values  
Safe and sustainable

## ABSTRACT

The synthesis of silver nanoparticles with controlled physicochemical properties is essential for governing their intended functionalities and safety profiles. However, synthesis process involves multiple parameters that could influence the resulting properties. This challenge could be addressed with the development of predictive models that forecast endpoints based on key synthesis parameters. In this study, we manually extracted synthesis-related data from the literature and leveraged various machine learning algorithms. Data extraction included parameters such as reactant concentrations, experimental conditions, as well as physicochemical properties. The antibacterial efficiencies and toxicological profiles of the synthesized nanoparticles were also extracted. In a second step, based on data completeness, we employed regression algorithms to establish relationships between synthesis parameters and desired endpoints and to build predictive models. The models for core size and antibacterial efficiency were trained and validated using a cross-validation approach. Finally, the features' impact was evaluated via Shapley values to provide insights into the contribution of features to the predictions. Factors such as synthesis duration, scale of synthesis and the choice of capping agents emerged as the most significant predictors. This study demonstrated the potential of machine learning to aid in the rational design of synthesis process and paves the way for the safe-by-design principles development by providing insights into the optimization of the synthesis process to achieve the desired properties. Finally, this study provides a valuable dataset compiled from literature sources with significant time and effort from multiple researchers. Access to such datasets notably aids computational advances in the field of nanotechnology.

## 1. Introduction

Already available on the market, silver nanoparticles (AgNPs) embedded into products offer a variety of benefits such as an antimicrobial functionalities by eradicating or inhibiting the growth of bacteria [1], fungi [2], and viruses [3], as well as the ability to prolong the shelf life of products (i.e., food packaging). Their versatility renders them valuable in various products' applications including wound dressings [4], medical devices [5,6], water purification systems [7], and antibacterial coatings [8]. AgNPs are also utilised as drug delivery carriers due to their size which enable them to readily penetrate cells and deliver therapeutic substances to targeted sites, enhancing treatments' efficacy [9]. Additionally, because their unique optical and electrical properties, they are utilized in sensing and detection systems [10],

measuring small quantities of target molecules, including heavy metals [11], environmental pollutants and biomarkers [12]. This has implications in fields like environmental monitoring, food safety and medical diagnostics [13]. They are also helpful in applications such as imaging devices [14] and solar cells [15]. Finally, have the ability to function as catalysts in a variety of industrial processes, such as organic synthesis, pollution control, and chemical manufacturing [16–18]. AgNPs have attracted attention as evidenced by the significant demand for and investment in associated research. The demand for AgNPs has been gradually increasing over the past 15 years, and manufacturing of the product is expected to exceed 500 tons annually to meet the demands of various industries [19]. Due to the growth of the market worldwide and the current offer of products with incorporated AgNPs, the study of their biological activity, functionalities and safety has become a paramount

\* Corresponding authors.

E-mail addresses: [lara.faccani@issmc.cnr.it](mailto:lara.faccani@issmc.cnr.it) (L. Faccani), [ilaria.zanoni@issmc.cnr.it](mailto:ilaria.zanoni@issmc.cnr.it) (I. Zanoni).

<sup>1</sup> This author shares dual first authorship with Irini Furxhi.

<https://doi.org/10.1016/j.csbj.2024.02.010>

Received 12 December 2023; Received in revised form 12 February 2024; Accepted 14 February 2024

Available online 17 February 2024

2001-0370/© 2024 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

matter of research [20]. However, their activities are modulated by their physicochemical properties (pchem) which are defined in the initial life cycle stage, the synthesis. To exploit and fulfil the Safe by Design (SbD) concept at the earliest stages of innovation it is imperative to comprehend which pchem properties, which are indirectly defined by key synthesis parameters (temperature, precursors, concentrations, treatments, etc), impact the toxicological and/or antibacterial profile. This challenge can be addressed with the development of predictive models that forecast desired endpoints while capturing the relationship among pchem and key synthesis parameters. From the standpoint of computational-based drug development approach, it is crucial to create predictive models that can screen a variety of potential experimental conditions and help identify the selection of optimal ones to employ in the laboratory experiments. In order to create such models, three key parts with their related data have to be investigated, i) synthesis methods of AgNPs along with experimental conditions, ii) resulting measured pchem properties of the synthesised AgNPs and iii) their functional or safety profiles.

### 1.1. Synthesis methods & experimental conditions

Top-down and bottom-up methods by either chemical, physical or biological means are the main synthesis approaches for AgNPs. A top-down example includes the grinding of bulk metals and the subsequent inclusion of colloidal protective agents to stabilize the resulting nanosized metal particles. The bottom-up approach includes reduction of metal precursors, electrochemical or green synthesis methods [21,22] etc., with the main ones being: i. chemical reduction of a silver salt such as silver nitrate ( $\text{AgNO}_3$ ) using a reducing agent (sodium borohydride ( $\text{NaBH}_4$ ) or sodium citrate ( $\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$ )) [23]. ii. sol-gel technique, which combines a gel-forming agent solution with a silver precursor (i.e., silica precursor). After that, the mixture is exposed to condensation and hydrolysis, which forms a gel containing silver ions. To convert ions into AgNPs, the gel is subjected to further heat treatment or reduction [24]. iii) green synthesis, which involves the use of natural or biological materials as reducing agents or stabilizers (i.e., plant extracts) [25]. iv) electrochemical deposition which involves the reduction of silver ions present in an electrolyte solution onto an electrode surface, typically a cathode [26]. During the synthesis process, several parameters can be adjusted to control and fine-tune the resulting pchem properties of AgNPs, such as:

- 1) the reagents' concentration which includes the silver precursor or the reducing agents. These chemical agents have an impact on the size, shape and dispersity index of the resulting NPs; higher concentrations frequently resulting in smaller NPs [27].
- 2) the pH of the reaction solution. The stability, aggregation behaviour, and interaction of the particles with other substances can all be affected by changes in the surface charge of the particles caused by the pH of the solution [28].
- 3) the reaction temperature. Temperature plays a crucial role in regulating the growth and nucleation kinetics; Higher temperatures often result in faster nucleation and growth, leading to larger NPs [29].
- 4) the usage of stabilizing agents (i.e. polymers or surfactants). Those agents can control stability and prevent aggregation, through their effects on steric hindrance and/or electrostatic repulsion, which can also affect surface chemistry and dispersibility [30].
- 5) the reaction time. Time has a strong impact on the pchem properties with longer reaction times often resulting to larger particles due to continued growth and reprecipitation processes associated with the Ostwald ripening [31,32].
- 6) the use of external fields such as electric, magnetic or ultrasound. AgNPs with desired size and morphology can be produced by adjusting the applied voltage and deposition duration [33,34].

- 7) the use of natural or biological materials as reducing agents or stabilizers in green synthesis methods can introduce unique properties due to synergistic interactions [35].

Based on the information summarized above, it can be hypothesized that various experimental conditions can be adjusted to alter the size, shape, and other pchem properties of synthesised NPs.

### 1.2. Physicochemical properties

One of the main pchem properties affecting NPs' behaviour is their core size. AgNPs are typically employed in products' applications ranging from 1 to 10 nm in size [36]. Numerous studies have revealed that smaller particles exhibit greater antibacterial activity but higher cytotoxic responses [37,38]. As the particle size gets smaller the specific surface area increases, hence, a greater proportion of its atoms are displayed on the surface [39]. This implies that biological interactions and toxicity are dependent on the particles' surface area than on the particle mass. The dispersibility, solubility, and/or hydrophobicity can be altered by the AgNPs' functionalization with various coatings. The coatings in turn, affect AgNPs' bioaccessibility and biodurability. There is a considerable literature amount indicates that the type of coating impacts both fate and toxicity of AgNPs [40–42]. For instance, trisodium citrate (CT-AgNP) and polyvinylpyrrolidone (PVP-AgNP) improve biocompatibility and stability against agglomeration phenomenon [43]. Moreover, coating modifies the surface charge which influences their cytotoxic potential and biological targets; positively charged NPs have stronger affinity for negative bacterial membranes causing higher cellular contents leakage and bacterial death [44]. Studies employing AgNPs of different sizes also revealed shape-dependent effects. For example, compared to spheres and wires, 55 nm silver nanotubes had a stronger antibacterial activity against *Escherichia coli* [45]. On the other hand, spherical, triangular and cuboid AgNPs were found to have no biocidal effect on *Staphylococcus aureus* [46]. The agglomeration state of NPs also determines cellular uptake as well as biological reactions [47]. A number of studies have shown that agglomerated particles exhibit reduced cytotoxicity compared to free AgNPs [47,48]. In summary, knowledge of the pchem properties is essential to comprehending the biological reactions of AgNPs and the underlying mechanisms of action which determined their applications.

### 1.3. AgNPs mechanisms of action and applications

AgNPs stand out among the many NMs that are known to be useful in preventing the growth of several bacteria due to their potent inhibitory and bactericidal properties [49,50]. AgNPs are also familiar for their antifungal and antiviral activities [51]. Several studies have reported their potent antifungal activity against several phytopathogenic fungi (e.g. *Alternaria alternata*, *Sclerotinia sclerotiorum*, etc) as well as human pathogenic fungi (e.g. *Candida* and *Trichoderma sp.*) [52] and against several viruses including SARS-CoV-2 [53,54]. Research has indicated that AgNPs interact with bacterial membranes but the precise mechanism underlying these activities remains unclear [55,56]. One leading hypothesis holds that AgNPs when in contact with bacteria, produce free radicals and reactive oxygen species, which damage the bacteria's internal organelles and change the intracellular signalling pathways that trigger apoptosis [57]. Another mechanism is the adhesion of AgNPs to the bacterial wall, followed by the infiltration of particles that act as a Trojan horse and release ions, causing bacterial membrane damage and leakage of cellular contents and death [58]. For example, the antibacterial activity of 12 nm AgNPs was excellent against both Gram-positive and Gram-negative bacteria including *Staphylococcus bacillus*, *Staphylococcus aureus*, and *Pseudomonas aeruginosa* [59]. This indicates that both membrane thickness and surface charge facilitate particle attachment onto the membrane [60]. Finally, the release of  $\text{Ag}^+$  ions from the surface area is another crucial factor that contributes to the antibacterial

activity. Release rate is dependent on a number of factors including size, shape, capping agent and colloidal state of NPs [51,61]. It is well established, smaller AgNPs have faster  $\text{Ag}^+$  release rate and, hence, an increased bioavailability, and also cytotoxicity, compared with larger NPs [62,63].

It becomes evident from the above sections that a modelling approach is essential to effectively and systematically capture the transversal information involved in the synthesis stage and to consider the impact of the many experimental parameters involved in the synthesis process on the pchem properties of the resulting AgNPs, their antibacterial efficiencies, and/or their toxicity effects.

#### 1.4. Machine learning models

With conventional approaches it is challenging to systematically explore the entire parameter space and comprehend the relationships between experimental synthesis conditions and resulting AgNPs characteristics due to the complexity and multidimensionality. Additionally, manual experimentation can be time-consuming, costly, and limited in scope, making it difficult to optimize synthesis conditions effectively. Hence, modelling techniques, such as machine learning (ML) models provide a promising means to navigate this complexity and capture the interplay of multiple parameters.

ML models can be trained using existing data on synthesis experimental parameters and corresponding pchem properties or functionalities, allowing the development of predictive models that can guide and optimize the synthesis process [64,65]. The models can then be used to predict the AgNPs' properties synthesized under different conditions. For example, by inputting the reagent concentrations, reaction temperatures, reaction duration etc., into a model, it can predict the resulting size, shape, or stability [66,67]. ML can be used to optimize the synthesis process by iteratively adjusting and identifying the optimal combination of synthesis parameters to converge into the desired properties. Designing effective experiments (Design of Experiments), to investigate the synthesis parameter space, can be achieved by ML which recommends the most instructive experiments to be conducted based on the knowledge gained from the available data. Such as approach minimizes the number of experiments required and speeds up the process of discovering the combination of synthesis conditions. Nanoparticles have only partially captured the benefits of automation, as exhaustively discussed in a review on the role of ML in Nanosafety [68]. One notable example of ML application is a proposed framework that optimizes the synthesis and extracts knowledge of the relationship between composition and optical properties [69]. The neural networks' output layer is composed of 421 nodes, which are corresponding to the UV-Vis spectral data points. The experimental input parameters contained variables of different flow rates such as silver nitrate, trisodium citrate etc., The framework was built based on flow rates as input features and specific experimentation apparatus to obtain synthesized in aqueous sub-microliter droplets NPs. In an attempt to increase the application domain of ML in the synthesis process, we manually collected synthesis data from multiple literature studies and leveraged various ML algorithms to explore the parameter space of synthesis conditions during the experiments and the resulting AgNPs characteristics followed by a targeted functionality (such as antibacterial properties). Data extraction included parameters such as reagent concentrations, reaction conditions, typology of stabilizing agents, etc. Additionally, pchem properties such as the core size measurements obtained from various characterization techniques, such as transmission electron microscopy (TEM), scanning electron microscopy (SEM), were recorded. Antibacterial efficiency and toxicological profiles were also collected when available in the study design. In a second step, based on data completeness, we employed regression ML algorithms to build predictive models targeting various endpoints to establish relationships between synthesis parameters and desired outputs. The developed models were validated using a cross-validation approach. Finally, the feature's impact, based on

SHapley Additive exPlanations (SHAP) average values, was utilized to provide insights into the contribution of individual features in the model's predictions.

## 2. Materials and methods

### 2.1. Literature review

A comprehensive literature review was performed for the duration of 2004–2022 using a combination of keywords such as “silver nanoparticles”, “silver nanomaterials”, “synthesis”, “nanoparticle synthesis”, “nanoparticle”, “chemical reduction”, and “green synthesis” to identify relevant studies and extract key information. The search was performed across various academic databases, including PubMed, Scopus and Google Scholar. The aim was to compile a variety of synthesis protocols including the experimental conditions and the steps followed, the characterization techniques employed for pchem properties identification and to record functionalities or toxicological properties of the aforementioned synthesized NPs. The retrieved articles were screened based on their titles and abstracts to assess their relevance to AgNPs synthesis. Selected articles were then obtained in full text and further reviewed to determine their suitability for inclusion in the literature review. Additionally, the reference lists of these articles were examined to identify additional relevant studies that may have been missed in the initial search. The inclusion criteria encompassed a variety of publication types, including research articles and review papers, excluding book chapters and conference proceedings.

### 2.2. Data extraction – first dataset

A manual data extraction was performed systematically from the selected articles. Relevant information, including synthesis parameters such as experimental conditions, AgNPs pchem properties, characterization techniques, toxicological outcomes, and antibacterial efficiency measurements were recorded in an Excel workbook format. To ensure consistency and reduce potential biases from expert judgment, the literature review was conducted by a team of researchers with expertise in nanomaterial synthesis and characterization. Regular discussions were held to address any discrepancies or uncertainties during the data extraction process. Having data on the synthesis procedure, the first step of the innovation process, is crucial for developing computational models for predicting multiple endpoints such as resulting pchem properties or functionalities. The data extraction included synthesis parameters such as:

- Reagent concentrations and the chemical compounds. As reagents, we considered the silver precursor, the reducing agents used, the capping/stabilizing agent, and any possible additional catalyst during the reaction process.
- Reaction conditions included temperature, any external source of energy applied (ultrasound, pressure etc.), the pH of the reaction, reaction duration, stirring approach and stirring conditions, and any specific environmental conditions (e.g., light conditions).
- Post-synthesis treatments were recorded such as drying or washing.
- Resulting pchem properties of the synthesized AgNPs were recorded. The characterization techniques employed to analyse the synthesized nanoparticles were specified such as transmission electron microscopy (TEM), scanning electron microscopy (SEM), X-ray diffraction (XRD), dynamic light scattering (DLS), etc., providing information about AgNPs size, shape, zeta potential etc.
- Finally, antibacterial and toxicological results were also extracted, reporting all the experimental details such as the cell line or organisms exposed, the exposure conditions (dose and time) and the endpoints measured along with the assays such as cell viability, inflammation indication (in case of toxicological studies) or the

antibacterial efficiency or zone of inhibition (in case of antibacterial assessments).

### 2.3. Data harmonisation and exploration

Following data extraction, the features were then harmonized to create a unified and cohesive dataset, to ensure compatibility and consistency, and to allow the exploration of features and the development of ML models. Experimental data collected from different studies exhibit variations in terms of synthesis parameters, characterization techniques, units of measurement, and reported properties. During meetings and discussions, the team collaborated to harmonize the extracted data and to reach a consensus. It was crucial to ensure that the dataset covered a broad range of synthesis conditions while maintaining data integrity. To achieve harmonization, the following main steps were followed as a principle:

- converting various units of measurements into a consistent format (SI unit). For example, ensuring that reagent concentrations are expressed in the same units (e.g., molar concentration) or that temperature is consistently reported in a specific scale (e.g., Celsius or Kelvin). This step minimized discrepancies arising from different reporting conventions.
- aligning the different terminologies used to express the same ontology. For example, synthesis parameters, such as reagent names, reducing agents, stabilizing agents, and experimental conditions, needed to be harmonized to ensure consistency across the dataset. This involved creating a standardized vocabulary to facilitate data integration.
- integration of features into simpler categories to avoid high cardinality features for the training of the models, for example, the capping agents might greatly vary across the studies for this reason, agents were simplified, classified, and aggregated into organic and inorganic classes (more information in the results Section 3.2).

Finally, all the data were inserted in a final workbook (*Raw data* tab in the [supplementary material](#)).

### 2.4. Machine learning exploration

The final workbook was split into i) synthesis dataset, ii) toxicological dataset, and iii) antibacterial dataset to explore the potential of developing and applying modelling tools based on data completeness. Data completeness in this work refers solely to the extent to which data required for modelling is present within the dataset with low data completeness suggesting missing information that hinders the application of modelling. In the nanosafety field data completeness can also refer to minimum reporting standards or guidelines for specific experiments or applications (e.g. OECD, ISO, REACH, MIRIBEL). When data sufficient for modelling exploration the selected dataset was then pre-processed via one-hot encoding for the categorical variables enabling the inclusion of these in predictive regression modelling tasks. One-hot encoding is a commonly used technique for handling categorical values converting them into a numeric label. In a second step, PyCaret regression library [70] was utilized to train various ML algorithms. PyCaret contains more than 18 regression models such as linear regression, decision tree, random forest, support vector, extra trees, ridge, elastic net, huber, bayesian ridge among others. These algorithms cover a wide range of popular regression techniques, enabling the comparison of different models. The models were trained with the training set and validated via 10-fold cross validation approach. In this method, the dataset is divided into ten subsets of equal size. The model is then trained on nine of these subsets (training set) and evaluated on the remaining subset (test set). This process is repeated ten times, with each subset serving as the validation test set once. By averaging the performance metrics across these iterations, such as accuracy or mean squared

error, the cross-validation approach provides an estimation of the model's performance on unseen data. This estimation accounts for variability in the data and helps assess the generalization ability and predictability of the model. Validation metrics such as root mean squared error (RMSE), mean absolute error (MAE), and R-squared ( $R^2$ ), were used to assess and compare the models' accuracy. In a final step, the features' impact based on SHAP (SHapley Additive exPlanations) average values [71,72] was derived to provide insights into the overall contribution of individual features in the model's predictions. A positive average impact value suggests that the feature leads to more accurate predictions to the model's output, while a negative value indicates the opposite. Features with higher average impact values have a stronger influence on the model's predictions, indicating their significance in determining the target variable. The feature SHAP average impact values help identify key drivers or variables that significantly contribute to the model's output and gain insights into the underlying relationships between the features and the output. The code for the ML pipeline containing the preprocessing steps and models exploration can be found here.<sup>2</sup>

## 3. Results

### 3.1. Literature review & data extraction

From the literature review, 219 studies were collected from the 2004–2022 period. In the first workbook, both synthesis and potential outcomes, covering antibacterial and toxicological assessments, were extracted. The process of including data into the workbook file involved duplicating each row based on a specific parameter, while keeping the remaining columns unchanged. For example, if the same synthesis process and resulting pchem properties were exposed into a cell culture under different exposure doses (i.e., 10, 50, 100 mg/ML), this resulted in three rows with only exposure dose being changed (same process for the synthesis methodology or the antimicrobial studies). This technique allowed for the systematic generation of multiple rows, each representing a unique combination of the duplicated parameter and the constant values of the other columns. By expanding rows in this manner, the dataset encompass various scenarios and conditions related to the parameter of interest, while maintaining consistency in the other attributes. This approach enables comprehensive analysis and facilitates further exploration of the data by considering the impact of different values of the duplicated parameter on the corresponding outcomes or observations.

Below, in Table 1 a brief overview of the targeted information extracted:

### 3.2. Data harmonization and split

Data extraction was followed by data harmonization, where the features were unified to create a cohesive dataset. Compatibility and cohesion are essential for ML model development and to enable feature exploration for insights from multiple studies we harmonized all units measurements for example minutes converted to hours for synthesis time and exposure duration (h), millilitre for volume (ML), millimolar for concentration (mM), Celsius degree for temperature ( $^{\circ}$ C), watt for external power (W), pound per square inch for pressure (psi), kilohertz for frequency (kHz), and milligram per millilitre for exposure dose (mg/ML), where feasible. In some cases units harmonization was not possible such as concentration expressed in either mass or volume metrics (i.e., nanoparticles count or NPs per milliliter (NPs/ML) where not possible to convert into a single metric.

Concerning the categorical features, we harmonized them in distinct classes. For example,

<sup>2</sup> <https://github.com/IrieIrie/Design-rules-AgNPs>

**Table 1**

Overall overview of the first workbook design created by an iterative discussion with the research group.

Synthesis studies			Physicochemical properties characterization			
<b>The synthesis method</b>	Synthesis process	Chemical reduction, sol-gel, green synthesis, phytochemical or others	<b>Pchem properties</b>	Chemical composition	The chemical composition followed by the method of determination and instrumentation used	
	Steps involved	Single-step or multiple steps process		Dissolution	Concentration of silver cations dissolved in relevant media ( $\mu\text{M}$ )	
	Synthesis duration	The duration of the entire synthesis process (h)		Shape	Shape followed by method of determination and instrumentation used	
<b>Precursor</b>	Scale	The scale of the synthesis process (ML)		Crystal structure	Crystal structure followed by method of determination and instrumentation used	
	Precursor	The precursor compound utilised (usually $\text{AgNO}_3$ )		Zeta potential	Zeta potential followed by method of determination and instrumentation used. In addition, standard deviation, pH and isoelectric point are reported (mV)	
	Source	The precursor commercial source (Sigma Aldrich, Aladdin, ACE chemicals, etc)		Core size	Core size followed by standard deviation, method of determination and instrumentation used (nm)	
	Purity	The precursor purity (%)		Crystallite size	Crystallite size followed by standard deviation, method of determination and instrumentation used (nm)	
	Concentration	The precursor concentration along with its units of measurement (mM, where not explicit or calculable we express the amount of precursor in g)		Hydrodynamic diameter	Hydrodynamic diameter followed by standard deviation, method of determination and instrumentation used (nm)	
	<b>Capping agent</b>	Compound		The capping agent used during synthesis		Surface area
Classification		Capping agent classification into plant-based, chemical, fungal extract, bacterial-based, etc	Polydispersity index	Polydispersity index followed by standard deviation, method of determination and instrumentation used		
Class		Organic versus inorganic agents	Agglomeration	Information if agglomeration was present		
Concentration		The capping agent concentration along with its units of measurement (mM or mg/ML)	UV-Vis peaks	The peaks of UV-Vis analysis for NPs and of reducing agents (nm)		
<b>Reducing agent</b>	Compound	The reducing agent used during synthesis		<b>Antimicrobial studies followed</b>		
	Classification	Reducing agent classification into plant-based, chemical, fungal extract, bacterial-based, etc		<b>Exposure conditions</b>	Exposure dose	The exposure concentration along with its units of measurement (mg/ML)
	Class	Organic versus inorganic agents		Exposure duration	The exposure duration of the experiment (h)	
<b>Catalyst</b>	Compound	The catalyst used during synthesis		Temperature	Temperature for the experimental conditions ( $^{\circ}\text{C}$ )	
	Concentration	The catalyst concentration along with its units of measurement (ML or mg/ML)		<b>Organisms</b>	Culture medium	The bacterial exposure medium
<b>Experimental conditions</b>	Temperature	Synthesis prevailing temperature. In case of a range of temperatures, the median value is reported ( $^{\circ}\text{C}$ )		Species	The species exposed	
	Order of reagent	The order of reagents in the synthesis process		<b>Outcome</b>	Strand	The specific strand of the organism
	Conditions of light	Dark, light or sunlight exposure			Bacteria reduction	The bacteria reduction along with its units of measurement (i.e., reduction in mm) and standard deviation
<b>External energy</b>	Energy	Addition of energy (through heating, sonication, ice bath, incubator, etc) or no addition of external energy		Bacterial methods	The methodology followed for the evaluation of the antibacterial efficiency	
	Power	In case of microwave or sonication, the power used (W)		<b>Toxicological studies followed</b>		
	Stirring	In case of stirring applied further split into magnetic, mechanic or shaking, etc		<b>Exposure conditions</b>	Exposure dose	The exposure concentration along with its units of measurement (mg/ML)
	Stirring Speed	The rpm of stirring applied		Exposure duration	The exposure duration of the experiment (h)	
	Pressure autoclave	The amount of pressure applied (psi)		<b>Cell culture information</b>	Cell line	Cell line name
<b>Post treatment</b>	Ultrasound	In case of sonication, the frequency used (kHz)		Cell type	Cancer or normal and the specific type (i.e., glioma, macrophages, vascular, etc)	
	Treatment	Any post treatment of AgNPs such as drying, washing, centrifugation, etc		Cell origin	The origin of the cell line (human, animal, plant).	
	Solution Methods	The solution where the treatment took place		Cell organ	The organ which the cell line represents	
	Speed	The methods applied for the specific treatment recorded above		Multiwell	The number of wells in the study	
	Duration	In case of centrifugation applied, the rpm speed		<b>Outcome</b>	Toxicological assay	Assay used to assess the toxicological endpoint
	The duration of the speed method (min)		Toxicological endpoint	The endpoint measured by the assay		

(continued on next page)

Table 1 (continued)

Synthesis studies		Physicochemical properties characterization	
Temperature	In case of drying process followed, the temperature applied (°C)	Toxicological results	The endpoint results are followed by the standard deviation and its units of measurements
Duration	The duration of the drying method (min)	Information of the studies such as the author, title, year, DOI are recorded. An initial notebook of 6202 rows was created.	
Other	Information such as the storage conditions		
Yield	The yield of synthesis reaction (%)		

- the synthesis process was divided into biosynthesis, micosynthesis, green synthesis, or general wet chemical synthesis.
- the capping agent feature was first classified into aggregated classes, for example *algae extract* aggregated class contains species such as: *Padina sp. marine* alga extract, *Noctiluca scintillans*, *Ecklonia cava*, etc, or the *fungus extract* aggregated class contains extract from *Aspergillus flavus*, *Emericella nidulans*, *Macrophomina phaseolina*, etc. In a second step, features were engineered to capture information on organic versus inorganic capping agent: organic contains information such as plant, fungal, fruit extracts, biopolymer or bacteria. The same approach was followed for the reducing agent.
- The order of reagents was converted into a codified feature where the silver precursor is denoted with A; reducing agent with B; capping agent with C; catalyst with D or other to simplify its representation. CADB for example would signify the following order of reagents: *capping;precursor;catalyst;reducing*. In that manner, the steps of synthesis can be represented by a single feature.
- The data was then split into synthesis dataset, pchem dataset, toxicological dataset and antimicrobial dataset for ML purposes.

### 3.2.1. Synthesis & targeting core size

For the synthesis dataset, briefly, from the nearly 900 rows of information (from 219 studies), phytosynthesis was the most common synthesis process reported (57.7 % of rows, 119 studies) followed by wet chemistry (26 % of rows, 58 studies) and biosynthesis (12.7 % of rows, 28 studies). Missing values in the dataset containing all the synthesis information and pchem information (no toxicological or antibacterial information) (Supplementary material, *synthesis tab*) were common. Roughly 89 % of the rows had single-step synthesis process (185 studies). When external energy was applied, heating was one of the most common (41 %, 93 studies) followed by sonication (10 studies) and microwave assistance (5 studies). Regarding the temperature applied during synthesis, nearly half of the values fall within the range 24–64 °C with seven studies applying > 100 °C. Regarding the light conditions, 86 % of the rows contained no information (we hypothesised that no special light conditions were applied). 29 studies reported dark conditions. The exact order of reagents was missing in 35 studies. The duration of synthesis was absent nearly 18 % (42 studies) of the rows, while the scale of synthesis (ML), 30 %. Features that were missing more than 80 % contained the precursor purity information. However, precursors' concentration (expressed in mM) had few missing values, same as capping agent compound (mostly plant or chemical agents) and reducing agents. On the other hand, capping agent and reducing agent concentrations (mg/ML) had nearly 40 % missing values. The pH during synthesis was missing > 70 % of the row cases. Regarding the pchem properties, dissolution was missing the majority of the time (> 90 %), shape nearly 30 %, crystal structure > 65 %, zeta potential and hydrodynamic size measurements > 70 %, while crystallite size, polydispersity index and surface area were almost always missing. Core size had the lowest count of missing values, nearly 25 % (6 studies out of 219 did not report core size). Since core size had the most data completeness and is one of the most crucial factors affecting biological interactions, antibacterial capabilities, and toxicity, was selected as model output to be predicted from synthesis parameters inputs.

3.2.1.1. Targeting core size. Rows with no core size data were dropped,

and when information on the standard deviation or size distributions on the core size was available (92 studies provided SD), those rows were triplicated containing the average, the minimum and maximum core size. This resulted in a new dataset (Supplementary material, *core size tab*) containing 2113 rows with 17 input features (7 numerical and 9 categorical) showed in Table 2. The stirring speed, capping and reducing agent concentrations features were excluded from the modelling part due to substantial amount of missing values.

Table 2

Input features used to predict core size along with the missing values presented in that dataset and a description of the features.

Group of features	Input features	Missing values	Description (min – max, mean values), categorical examples
The synthesis method	Synthesis process	0.0 %	Such as phytosynthesis (59 %), micosynthesis (4.8 %), wet chemical (24.5 %), biosynthesis (11 %), etc
	Steps involved	0.0 %	Single step synthesis (85 %), or multiple steps
	Synthesis duration	12.8 %	0.008 – 336, 17.5 (h)
External energy	Scale	24.5 %	1 – 510, 87.2 (ML)
	Energy	0.0 %	Such as heating (46 %), sonication (5.2 %), microwave (1.2 %), ice bath, UV treatment (1.8 %), cooling, none, etc
Stirring	Stirring	0.0 %	Such as mechanic (17.3 %), normal stirring (15.2 %), magnetic (13 %), shaking (12.9 %), yes (when no additional information was provided, 1.9 %), etc
	Stirring Speed	49.0 %	0 – 5000, 71 (rpm)
Experimental conditions	Temperature	1.3 %	-4 – 200, 48.4 (°C). In case of no information in the article, room temperature 25 °C was inserted
	Precursor	1.0 %	0.1 – 1000, 6 (mM)
Capping agent	Concentration	1.0 %	Such as plant (55 %), chemical (14.5 %), biopolymer (8 %), fungal extract (5 %), biomolecule (4 %), animal derivative (1.8 %), algae extract, etc
	Classification	0 %	Organic (92 %), inorganic or none
	Class	0.0 %	0.005 – 1000, 45.7 (mg/ML)
Reducing agent	Concentration	41.0 %	Such as plant (56 %), chemical (16 %), biopolymer (5 %), biomolecule, animal derivative, fungi extract, etc
	Classification	5.3 %	Organic (83 %), inorganic
Experimental conditions	Concentration	35.9 %	0.0025 – 500, 16.1 (mg/ML)
	Order of reagent	17.5 %	Codified feature such as AB (44 %), CADB, BA (20 %), etc, where, A: precursors addition, B: reducing agent addition, C: capping agent, D: catalysts, E: other step
	Class	5.3 %	Organic (83 %), inorganic
Post treatment	Treatment	3.8 %	Such as washed (42 %), no post treatment (43 %), dried (2.2 %), centrifuged, pH adjusted, filtration, etc
Outcome to be predicted	Core size	0.0 %	0.1 – 334.8, 28.7 (nm)

In the next step, missing values were imputed via iterative ML based imputation. During each iteration, the imputation model uses the imputed values from previous iterations along with other available features to improve the accuracy of predictions. This process continues for a predefined number of iterations or until convergence is achieved. The dataset was randomly stratified split into ten training (1479 rows) and test sets (634 rows) for training and validation ML purposes. Models were validated via 10 fold cross validation approach which provides valuable insights into the predictive power of a model on unseen data from the training set.

### 3.2.2. Toxicological

In case of the toxicological dataset, the creation of a dataset targeting a toxicological outcome with synthesis parameters as inputs was not feasible (Supplementary material, toxicological tab). The diversity of outcomes did not allow the selection of a single outcome of sufficient size and data completeness. For example, various endpoints were measures such as cell viability, Inhibitory Concentration 50 % (IC50), Lethal concentration 50 % (LC50, LC90), minimum inhibitory concentration (MIC), lipid peroxidation, apoptosis, rate of haemolysis, scavenging activity, proliferation, etc. IC50 expressed as mg/ML, was the most commonly reported outcome in our case (20 studies), resulted only in nearly 500 rows of information. Thus, we did not proceed further with the toxicological case.

### 3.2.3. Antibacterial efficiency

From the 219 studies, 134 of them provided an antibacterial efficiency assessment using diverse endpoints such as biofilm inhibition/formation (8 studies, endpoint expressed as % or CFU/ML), LC50 (2 studies), growth rate, inhibitory concentration (IC10 or IC50, 4 studies), minimum bactericidal concentration (MBC, 6 studies), minimum inhibitory concentration (MIC, MIC50, MIC90, 28 studies with endpoint expressed as mg/ML, % or NPs/ML, mM, etc), zone of inhibition (ZOI, 88 studies), etc. Based on the analysis, ZOI (mm) was selected as the endpoint to be modelled which represents a circular area around the spot of the agent in which the bacteria colonies do not grow (expressed in mm) and indicates the effectiveness of the antimicrobial agent. The model captures the efficiency against multiple strains providing a comprehensive understanding of its overall effectiveness and generalizable findings. It may be more appropriate to develop separate predictive models for each bacterial strain for a more focused and accurate assessment of the relationship between antimicrobial agents and bacterial growth inhibition. However, due to limited availability of data for individual strains we combined information to assess the overall antimicrobial activity in a holistic perspective.

Rows absent with ZOI values were dropped, and when information on the standard deviation on the ZOI results were available (15 studies provided SD), those rows were tripled to contain the average, the minimum and maximum ZOI values. This resulted in a new dataset (1251 rows with 25 features, 9 numerical and 16 categorical) shown in Table 3 (Supplementary material, antimicrobial tab). Scale synthesis, capping and reducing agent concentrations and stirring speed were excluded from the modelling part due to substantial amount of missing values.

### 3.3. Machine learning exploration

PyCaret regression library [70] was utilized to train various ML algorithms (linear regression, decision tree regression, random forest regression, and support vector regression, among others). Models parameterization or optimization was not performed since the chosen algorithms have few hyperparameters and are relatively straightforward. In addition, the emphasis of the study is placed on the exploration of the feasibility of modelling approach using literature data rather than the parameter optimization. The models were trained using the training sets, which consisted of subsets of the available data (70 %). A

**Table 3**

Input features used to predict zone of inhibition along with the missing values presented in that dataset and a description of the features.

Group of features	Input features	Missing values	Description (min – max, mean values), categorical examples
The synthesis method	Synthesis process	0.0 %	Green synthesis, phytochemical synthesis (70 %), wet chemical (9 %), biosynthesis (8.2 %), etc
	Steps involved	0.0 %	Single step synthesis (95 %), or multiple
	Synthesis duration	11.3 %	0.008 – 168, 16.2 (h)
External energy	Scale	40.6 %	3 – 510, 107.2 (ML)
	Energy	0.0 %	Such as heating (31 %), autoclave (4.2 %), sonication (1 %), microwave (3.7 %), ice bath, UV treatment
	Stirring	0.0 %	Such as mechanic (18.3 %), normal stirring (28.7 %), magnetic (10.7 %), shaking (5.2 %), etc
Experimental conditions	Stirring Speed	49.0 %	0 – 5,000, 71 (rpm)
	Temperature	3.8 %	20 – 200, 44 (°C) in case of no information in the article, room temperature was inserted
Precursor Capping agent	Concentration	2.4 %	0.1 – 25,000, 19.2 (mM)
	Classification	0.0 %	Such as plant (64 %), chemical (8.7 %), biopolymer, biomolecule, fungal extract (7.3 %), algae extract (7.8 %), etc
Reducing agent	Class	1.0 %	Organic, inorganic
	Concentration	34.1 %	0.02 – 500, 68.6 (mg/ML)
	Classification	1.4 %	Such as plant (65.7 %), chemical (9.5 %), biopolymer, biomolecule, animal derivative, algae extract (7.8 %), etc
Experimental conditions	Class	1.4 %	Organic, inorganic
	Concentration	41.6 %	1.0 – 500, 36.4 (ML)
	Order of reagent	7.1 %	Codified feature such as ABE, CADB, CAB, etc, where, A: precursors addition, B: reducing agent addition, C: capping agent, D: catalysts, E: other step
Post treatment	Treatment	0.2 %	Such as none, (51.6 %), washed (26.6 %), dried (4.2 %), centrifuged, pH adjusted, filtration (3.8 %), etc
Pchem properties	Shape	9.2 %	Such as spherical (76.8 %), prism, cubic (4.8 %), hexagonal, triangular, etc
	Core size	9.4 %	0.68 – 334.8, 32.1 (nm)
	Core size (method)	10.3 %	Such as TEM (51.3 %), SEM (22.8 %), AFM (6 %), FETEM (5.7 %), FESEM (2.2 %), etc
Exposure conditions	UV-Vis peaks of NPs	6.1 %	280 – 780, 429 (nm)
	Exposure dose	16.2 %	0.00001 – 1000, 9.2 (mg/ML)
Organisms	Exposure duration	13.8 %	3 – 168, 26.8 (h)
	Culture medium	17.3 %	Such as Corkborers, Luria Bertani, Agar, Mueller Hinton, etc
	Species	0.0 %	Such as <i>E. coli</i> , <i>Staphylococcus</i> , <i>Bacillus</i> , <i>Klebsiella</i> , <i>Salmonella</i> , <i>Candida</i> , etc
Outcome to be predicted	Bacteria reduction (endpoint: zone of inhibition)	0.0 %	0 – 80, 12.7 (mm)

cross-validation approach to evaluate the models' robustness was used to validate models iteratively, with each fold serving as a validation unseen test set. This process was repeated ten times, ensuring that each fold was used also as the validation set. During cross-validation, various metrics to assess the accuracy and generalization capability of the models were assessed such as the root mean squared error (RMSE), mean absolute error (MAE), and R-squared ( $R^2$ ). The RMSE measures the average magnitude of the prediction errors and provides an indication of how well the model fits the data. The MAE represents the average absolute difference between the predicted and actual values and is useful for understanding the model's overall accuracy. The  $R^2$  score quantifies the proportion of variance in the target variable that is explained by the model and indicates the goodness of fit.

### 3.3.1. Core size prediction

The performance metrics of the 10-fold cross-validation results of the top five model's out of the 18 regression algorithms contained in the PyCaret library, shown in Table 4. The regression algorithms performed similarly with random forest and Extreme Gradient Boosting reaching during the 10-fold cross a  $R^2 \approx 0.70$ ,  $RMSE \approx 16.85$  and  $MAE \approx 9.10$  followed by Extra Trees Regressor.

Prediction error plot of random forest model (right) and residuals plot (left), of one individual instance of the ten subsets for visualization purposes, are shown in Fig. 1. The blue dots on residuals plot represent the training set instances where the model is trained on 70 % of the data and shows an  $R^2$  value of 0.84. The green dots represent the test set instances where the model is tested on 20 % of the data as a test set and shows an  $R^2$  value of 0.66. The residuals plot allows to identify areas of the target that may be more or less prone to errors by displaying the difference between residuals on the y-axis and the dependent variable on the x-axis. The plot indicates a good-performing model since the residuals are randomly scattered. In the right panel, the actual targets from the dataset are plotted against expected values produced by random forest model in a prediction error plot. Comparing this plot to the 45 degree line (identity line), where the prediction exactly matches the model, allows diagnosing the regression model fit. The plot shows a good fit of the model with an  $R^2$  value of 0.66 displaying a correlation between actual core size values and its predicted values.

We computed the mean absolute SHAP (SHapley Additive exPlanations) value as a measure to compare the influence of features in the model. The mean absolute SHAP value represents the average impact of a feature on the model's predictions across all instances. By comparing the mean absolute SHAP values across different features, we gained insights into their relative importance in predicting the target variable. Upon analyzing the results, we observed that certain features exhibited higher mean absolute SHAP values, indicating their stronger influence on the model's predictions (Fig. 2). These features collectively play a more significant role in determining the target variable and had a larger impact on the model's output. Conversely, features with lower mean absolute SHAP values contributed less to the predictions and had a relatively smaller influence.

- The synthesis duration showed the highest SHAP value indicating its strong influence on the prediction values of the core size. The duration plays a crucial role in determining the structural and

morphological characteristics of AgNPs; longer synthesis durations often result in larger particle sizes. The duration is closely linked to the kinetics and reaction pathways involved in the formation of NPs, due to a balance between growth and stabilization processes of NPs associated to reducing and capping agents [73]. Different synthesis durations can influence the rates of nucleation, growth, aggregation, or surface modification processes.

- The scale of synthesis followed second which could be explained by its function in determining the reagents interactions during synthesis. Changing the synthesis scale can affect the local ratios of reagents, which can influence the nucleation and growth kinetics, particle size distribution, and local composition [74]. It can also be attributed to its influence on the quantity- and composition-dependent NPs properties.
- The order of reagents was ranked third. The order can affect the nucleation and growth processes involved in synthesis [75]. Controlling the sequence of reagent addition, can influence the formation, capturing critical information about the kinetics and thermodynamics of nucleation and growth, and making it a significant predictor.
- The nature of the capping and reducing agent appeared high, since they provide information about AgNPs chemical nature that influences the surface chemistry and properties of the nanoparticles [76]. Consequently, the classification or the type of capping and reducing agent becomes a crucial predictor for modelling these nanoparticle characteristics, leading to high SHAP value. The reducing capping agent derived from plant is the most relevant in the top ten features, while the nature of the capping agent is followed in as biopolymer > chemical > plant.
- Stirring mode, concentration of precursor and type of post-synthesis treatment also appeared in the top ten SHAP values.

### 3.3.2. Zone of inhibition prediction

The performance metrics results of the top five model's of PyCaret's regression modules shown in Table 5. The regression algorithms performed similarly during the 10-fold cross validation reaching a  $R^2 \approx 0.72$ ,  $RMSE \approx 4$  and  $MAE \approx 2.5$  with Light Gradient Boosting Machine. Similar results were observed with random forest and extra trees regressor algorithm.

Prediction error plot (right) and of residuals error plot (left) of lightgbm regression model are shown in Fig. 3. The blue dots represent the training set instances where the model is trained on 70 % of the data and shows an  $R^2$  value of 0.87. The green dots represent the test set instances where the model is tested on 20 % of the data as a test set and shows an  $R^2$  value of 0.73. The plot shows fairly uniformly distributed residuals against the target in two dimensions indicating a good performing model. The prediction error plot shows a good fit of the model with an  $R^2$  value of 0.73 displaying a correlation between actual ZOI values and its predicted values.

We computed the mean absolute SHAP (SHapley Additive exPlanations) value as a measure to compare the influence of features in our model Fig. 4.

- Core size highest SHAP value in predicting the ZOI suggests its strong influence on the models' predictions. The model reflected what is known by mechanistic investigations. Smaller NPs with larger surface areas have increased contact area with bacterial cells, enhancing the chances of interactions. The core size affects the diffusion and penetration of NPs into bacterial cells. Smaller NPs can more readily penetrate the bacterial membrane, being internalized, accessing intracellular targets, and disrupting essential processes, causing cell damage [77].
- The core size can influence the targeting and specificity of NPs towards bacteria, which can explain the bacterial species features appearing second. For instance, certain core sizes may exhibit stronger interactions with specific bacterial strains or surface

**Table 4**

Top five machine learning regression algorithms validation results metrics (mean values of 10-fold cross validation approach).

Model	Acronym	MAE	RMSE	$R^2$
Random Forest Regressor	rf	9.11	16.82	0.70
Extreme Gradient Boosting	xgboost	9.06	16.89	0.70
Extra Trees Regressor	et	9.03	17.19	0.68
Decision Tree Regressor	dt	9.09	17.6	0.66
Light Gradient Boosting Machine	lightgbm	10.12	18.20	0.66



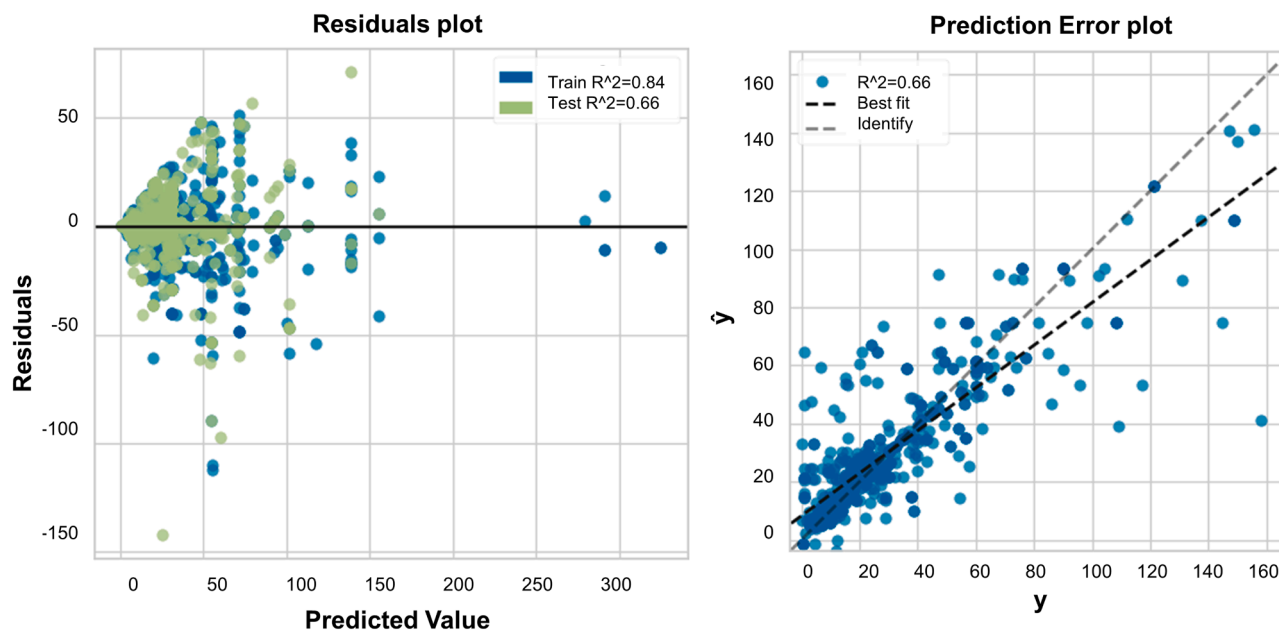


Fig. 1. Residuals plot (left) and prediction errors plot (right) for random forest model. Train set: 1479 rows, test set: 634 rows. The figure represents one instance of the ten fold validation process.

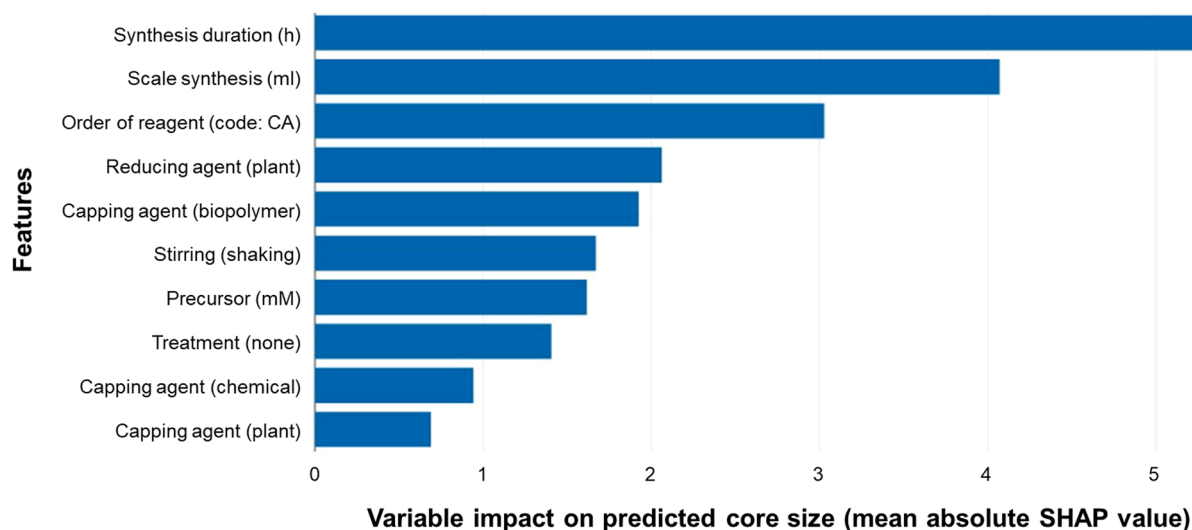


Fig. 2. Top 10 variables impact on predicted core size values (mean absolute SHAP value).

Table 5

Top five machine learning regression algorithms validation metrics (mean values) via 10-fold cross validation results.

Model	Acronym	MAE	RMSE	R <sup>2</sup>
Light Gradient Boosting Machine	lightgbm	2.49	4.07	0.72
Random Forest Regressor	rf	2.31	4.10	0.70
Extra Trees Regressor	et	2.27	4.20	0.68
Extreme Gradient Boosting	xgboost	2.35	4.432	0.65
K Neighbors Regressor	knn	3.02	4.71	0.61

receptors, leading to enhanced antibacterial activity [78,79]. Understanding the size-dependent targeting capabilities is crucial for predicting the antibacterial efficiency of NPs.

- Exposure dose appeared third in the SHAP, since the antibacterial activity of NPs often exhibits a concentration-dependent behavior. Higher exposure concentrations provide a higher concentration of

NPs (higher bioavailability), enabling more extensive interactions with bacterial cells and facilitating stronger inhibition of bacterial growth.

- The core size affects various properties of NPs, such as solubility, stability, and aggregation behavior. These properties can influence the dispersibility and behavior of NPs in the bacterial environment, impacting their antibacterial efficacy which can explain the type of culture medium appearing as top influential features (Muellet Hinton > Agar medium). For example, smaller NPs may exhibit higher stability and remain dispersed, ensuring sustained antibacterial activity [80,81].
- Synthesis process and parameters also appear in the top five, directly linked to the final size (as mentioned in a previous session).

Consequently, the core size, bacterial species, culture medium and synthesis process become critical predictors for modelling the antibacterial efficiency, leading to their high SHAP values. Synthesis duration,

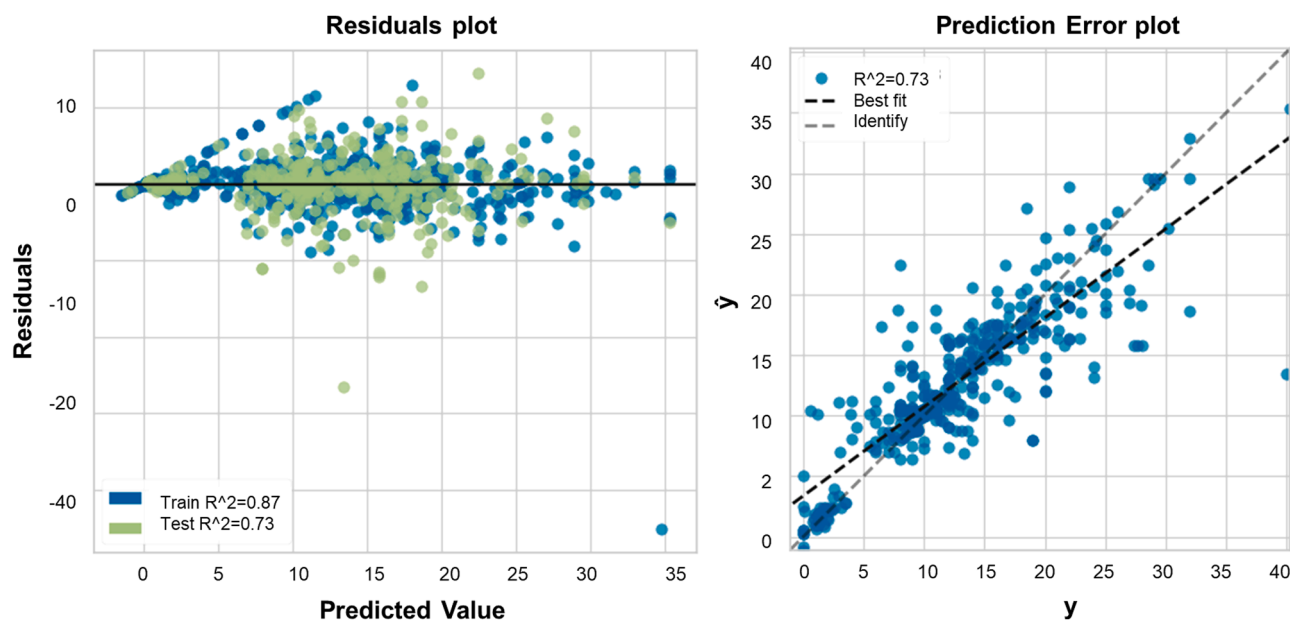


Fig. 3. Residuals plot (left) and prediction errors plot (right) for lightgbm regression model. Train set: 875 rows, test set: 376 rows.

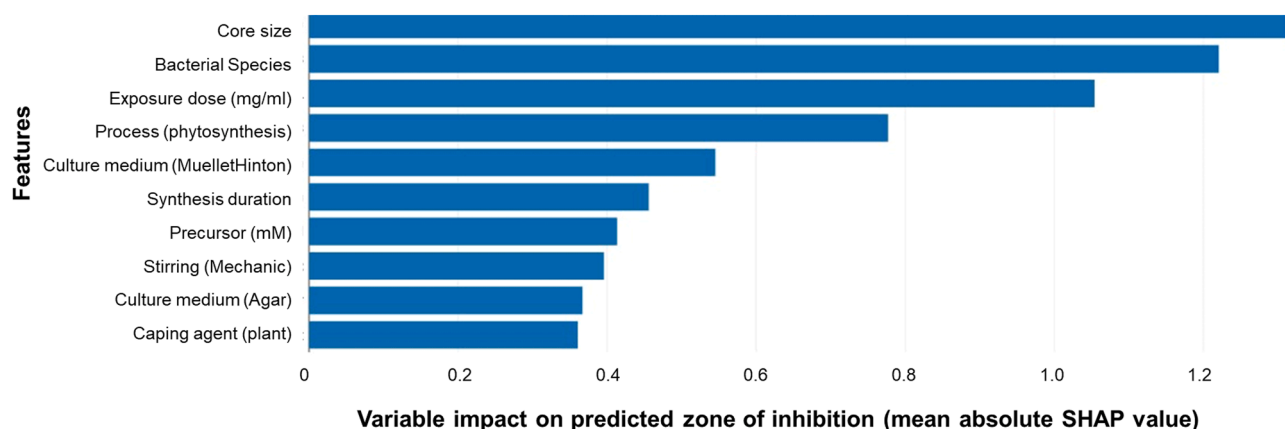


Fig. 4. Variable impact on predicted bacterial reduction (mean absolute SHAP value).

stirring and capping agent also appeared in the top ten SHAP values.

#### 4. Discussion

AgNPs are used in various products, such as antibacterial soaps, wound dressings, bandages, medical devices, certain cosmetics and personal care products (creams, lotions, deodorants, and shampoos) for their antimicrobial and antioxidant properties [82–85]. Those functionalities have been attributed to their unique pchem properties, thus manipulating those or predicting the properties during the synthesis stage is crucial for the final nano-enabled products. Overall, the ability to manipulate and/or predict pchem properties during the synthesis stage empowers researchers and manufacturers to create materials with tailored functionalities, optimized performance, and enhanced applicability in various industries. This approach also paves the way for innovative and advanced materials that can revolutionize technologies. By leveraging computational approaches, researcher can uncover synthesis parameters and pchem combinations that could enable the development of materials with tailored properties. In addition, predicting pchem properties during synthesis can lead to more efficient and cost-effective processes. By avoiding the need for trial and error in experimental synthesis, researchers can save time, resources, and

materials. AgNPs can be synthesized through various methods and the synthesis process can be complex, involving multiple parameters that influence the resulting NPs characteristics. Addressing this challenge necessitates the development of efficient predictive models that can capture the relationship among all the synthesis variables. In this study, a first attempt of utilizing synthesis data manually extracted from the literature with ML approaches has been reported. The data-driven approach has benefits but also carries some limitations.

##### 4.1. Benefits of a data driven approach

A modelling approach allows for the systematic investigation of AgNP synthesis by considering a wide range of parameters simultaneously. This holistic perspective facilitates a comprehensive understanding of the synthesis process, enabling the identification of optimal parameter combinations and conditions for desired AgNPs properties. ML, offers a valuable approach to be incorporated into the synthesis of AgNPs. It can effectively handle diverse datasets, capture nonlinear relationships, and make predictions [86,87]. By utilizing ML algorithms, researchers can consider multiple synthesis parameters and their interactions, enabling accurate predictions of AgNP properties based on given synthesis parameters, helping to guide and prioritize experimental

efforts [88]. It becomes possible to explore virtually the effects of different synthesis conditions combinations without conducting numerous time-consuming and costly experiments [89]. This speeds-up the synthesis process, reduces trial-and-error experimentation, and provides insights into the most influential parameters [90]. This approach promotes a more informed and data-driven approach to AgNP synthesis, leading to improved control over NPs characteristics and functionality. ML can aid in the discovery of novel AgNP formulations by guiding researchers towards unexplored regions of the parameter space [91]. This can lead to the identification of new NPs and functionalities that were not readily apparent through traditional synthesis approaches. The collaboration between traditional synthesis methods and ML-based approaches opens up new possibilities for innovation, efficiency, and impact in the field of materials science and nanotechnology. The feature importance analysis also provided valuable quantitative insights into the underlying factors that contribute significantly to the prediction of the outputs. By identifying them, deeper insights into the synthesis process and understand the parameters that have the strongest impact on the properties of AgNPs can be gained. This knowledge helps in streamlining the synthesis process, contributing to the development of more comprehensive and accurate models for predicting and optimizing AgNP synthesis. However, it is important to acknowledge limitations regarding data availability, generalizability, and assumptions made by the models.

#### 4.2. Limitations of the data driven approach

The ML models developed in this study are trained and validated using a specific dataset collected from the literature. The applicability and generalizability of these models to other synthesis conditions, NPs systems, or antibacterial agents should be approached with caution. Therefore, further validation and testing on independent datasets are necessary to assess the model's robustness and generalizability. The selection of features used in the models is critical for their predictive performance [92]. In this study, certain features were chosen based on their data completeness (missing values %). However, other important features may not be considered or available in the collected dataset. It is important to consider a broader range of features and explore their potential impact on the model's performance in future studies such as the yield of reaction and other pchem properties or atomistic descriptors. While ML models can provide insights into feature importance and their impact on predictions (e.g., SHAP values), they do not inherently provide causality for example high SHAP values associated with certain features indicate their influence on the model's predictions, but do not establish causal relationships. Therefore, caution should be exercised when interpreting the results and making causal claims based solely on the model outputs. For example, the synthesis duration is the most relevant feature to predict the final core size, followed by the synthesis scale and the order of reagents in the process [93–95]. Reducing and capping reagents' typology and concentration are less relevant in the final outcome compared to the three features mentioned above. Conducting experiments under controlled conditions using the predicted optimal synthesis parameters can help validate the model's predictions and assess their practical utility. Addressing these limitations and further exploring the predictive performance of the models in real-world scenarios will be crucial for the practical application and wider acceptance of ML approaches in the field of NPs synthesis and antibacterial research.

#### 4.3. The main limitation of data driven approaches: data

One of the main limitations of this study is the reliance on manually collected data from the literature. The availability and quality of the data might vary [96]. This lack of data poses challenges for researchers and hinders the development of accurate predictive models [97,98]. It becomes difficult to establish general guidelines or best practices for

achieving desired NPs characteristics. The accuracy and completeness of the data depend on the reporting and documentation of the synthesis procedures in the literature. Data harmonization is a crucial step in the development of ML models [99]. It involves standardizing and integrating diverse datasets from various sources to ensure compatibility and consistency, identifying common parameters, variables, and units of measurement across the studies. The lack of standardized reporting formats across different studies introduces variability in the collected data. Therefore, harmonizing the data is essential to create a unified and cohesive dataset for model development. It is crucial to collect comprehensive and diverse datasets that cover a wide range of synthesis conditions and nanoparticle properties to ensure the accuracy and generalizability of the models.

#### 4.4. Future directions

To overcome the scarcity of data, researchers can adopt strategies to generate and share synthesis data. Collaboration among researchers and institutions can facilitate the pooling of data from different studies and contribute to a collective knowledge base [100]. Additionally, efforts to publish detailed synthesis protocols and characterization data can help bridge the data gap and foster reproducibility in the field. Furthermore, emerging techniques such as high-throughput experimentation and automated synthesis platforms can aid in generating large volumes of data on AgNPs synthesis. These approaches can rapidly explore a wide range of synthesis conditions and provide valuable insights into the effects of various parameters on nanoparticle properties. Sharing such data openly can encourage collaboration, accelerate progress, and enable the development of machine learning models and predictive tools [89]. To achieve harmonization, standardizing the data involves converting various units of measurement into a consistent format. For example, ensuring that reagent concentrations are expressed in the same units (e.g., molar concentration) or that the same approach is used for NPs characterization. This step minimizes discrepancies arising from different reporting conventions. Aligning the parameters involves mapping and matching the terminologies used in different studies. Synthesis parameters, such as reagent names, reducing agents, stabilizing agents, and experimental conditions, need to be harmonized to ensure consistency across the dataset. This may involve creating a standardized vocabulary or ontology to facilitate data integration and fairness. Strong efforts are ongoing in the field of nanosafety towards the principles of Findability, Accessibility, Interoperability and Reusability (FAIR) data and their importance for modelling studies and data-driven safe and sustainable application of nano- and advanced materials [100–103].

Another emerging concept in the nanotechnology computational field is the use of read-across structure-activity relationship (RASAR) [69]. The integration of RASAR in the synthesis of AgNPs with ML models enhances the predictive capabilities and the efficiency of the models. RASAR leverages the similarity between chemical structures and their corresponding activities, enabling the extrapolation of properties to new or untested compounds. In this context, RASAR can be applied to complement the abovementioned models and enable the prediction of the pchem properties, antibacterial efficiencies, and toxicological profiles of AgNPs beyond the specific data points manually extracted from literature sources. RASAR extends the predictive scope by considering structural similarities with known compounds, providing a broader understanding of the relationships between synthesis parameters and endpoints.

## 5. Conclusions

This study employed a data-driven approach to predict the properties of AgNPs and their antibacterial efficiencies from synthesis parameters. The models were trained and validated using manually collected data from the literature, and their performance was evaluated using various

metrics. The feature importance analysis helped identify the influential variables in predicting the core size and zone of inhibition outputs. Factors such as synthesis duration, scale of synthesis, and the choice of capping agents were found to be important predictors in the models. The generalizability of the models to different datasets and experimental conditions should be further investigated. The development of standardized reporting formats for synthesis data would facilitate data collection and promote the reproducibility of studies in this field. Despite these limitations, the study showcases the potential of ML and their predictive capacity. The findings contribute to the growing body of knowledge in nanoscience and provide a foundation for further exploration and optimization of NPs synthesis processes by providing a data asset to be further explored and enriched. Ultimately, the integration of ML and experimental validation holds promise in advancing the development of materials, providing indication on how to implement a Safe-by-Design approach at the earliest stages of innovation.

### CRedit authorship contribution statement

Conceptualization, I.F., L.F., I.Z. and A.L.C.; methodology, I.F., L.F., I. Z., A.B. and M.V.; software, I.F.; validation, I.F., L.F., I.Z., A.B. and M.V.; formal analysis, I.F.; investigation, I.F., L.F., I.Z., A.B. and M.V.; resources, A.L.C.; data curation, L.F., I.Z., A.B. and M.V.; writing—original draft preparation, I.F., L.F., I.Z., A.B. and M.V.; writing—review and editing, I.F., L.F., I.Z., A.B., M.V. and A.L.C.; visualization, L.F.; supervision, A.L.C.; project administration, A.L.C.; funding acquisition, A.L.C. All authors have read and agreed to the published version of the manuscript.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

This work was supported by the “ASINA” (Anticipating Safety Issues at the Design Stage of Nano Product Development) project under European Union’s Horizon 2020 Research and Innovation Programme under grant agreement N. 862444, by the project “ECOSYSTEM FOR SUSTAINABLE TRANSITION IN EMILIA ROMAGNA” (ECS 0000003 - PNRR) Spoke 3 Green Manufacturing for Sustainable Economy, and by HORIZON MSCA-2022-PF-01-01 Programme, grant N°101103082.

### Data availability

The datasets for this study can be found as a supplementary material.

### Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.csbj.2024.02.010](https://doi.org/10.1016/j.csbj.2024.02.010).

### References

- [1] Abeer Mohammed AB, Abd Elhamid MM, Khalil MKM, Ali AS, Abbas RN. The potential activity of biosynthesized silver nanoparticles of *Pseudomonas aeruginosa* as an antibacterial agent against multidrug-resistant isolates from intensive care unit and anticancer agent. *Environ Sci Eur* 2022;34:1–15. <https://doi.org/10.1186/s12302-022-00684-2>.
- [2] Li L, Pan H, Deng L, Qian G, Wang Z, Li W, et al. The antifungal activity and mechanism of silver nanoparticles against four pathogens causing kiwifruit post-harvest rot. *Front Microbiol* 2022;13:1–16. <https://doi.org/10.3389/fmicb.2022.988633>.
- [3] Jeremiah SS, Miyakawa K, Morita T, Yamaoka Y, Ryo A. Potent antiviral effect of silver nanoparticles on SARS-CoV-2. *Biochem Biophys Res Commun* 2020;533:195–200. <https://doi.org/10.1016/j.bbrc.2020.09.018>.
- [4] Krishnan PD, Banas D, Durai RD, Kabanov D, Hosnedlova B, Kepinska M, et al. Silver nanomaterials for wound dressing applications. *Pharmaceutics* 2020;12:1–27. <https://doi.org/10.3390/pharmaceutics12090821>.
- [5] Marassi V, Di Cristo L, Smith SGJ, Ortelli S, Blossi M, Costa AL, et al. Silver nanoparticles as a medical device in healthcare settings: A five-step approach for candidate screening of coating agents. *R Soc Open Sci* 2018;5. <https://doi.org/10.1098/rsos.171113>.
- [6] Xu L, Wang YY, Huang J, Chen CY, Wang ZX, Xie H. Silver nanoparticles: Synthesis, medical applications and biosafety. *Theranostics* 2020;10:8996–9031. <https://doi.org/10.7150/thno.45413>.
- [7] Zahoor M, Nazir N, Iftikhar M, Naz S, Zekker I, Burlakovs J, et al. A review on silver nanoparticles: Classification, various methods of synthesis, and their potential roles in biomedical applications and water treatment. *Water (Switz)* 2021;13:1–28. <https://doi.org/10.3390/w13162216>.
- [8] Bruna T, Maldonado-Bravo F, Jara P, Caro N. Silver nanoparticles and their antibacterial applications. *Int J Mol Sci* 2021;22. <https://doi.org/10.3390/ijms22137202>.
- [9] Gomes HIO, Martins CSM, Prior JAV. Silver nanoparticles as carriers of anticancer drugs for efficient target treatment of cancer cells. *Nanomaterials* 2021;11. <https://doi.org/10.3390/nano11040964>.
- [10] Beck F, Loessl M, Baeumner AJ. Signaling strategies of silver nanoparticles in optical and electrochemical biosensors: considering their potential for the point-of-care. *Microchim Acta* 2023;190. <https://doi.org/10.1007/s00604-023-05666-6>.
- [11] Rossi A, Zannotti M, Cuccioloni M, Minicucci M, Petetta L, Angeletti M, et al. Silver nanoparticle-based sensor for the selective detection of nickel ions. *Nanomaterials* 2021;11:1–16. <https://doi.org/10.3390/nano11071733>.
- [12] Zahran N, Khalifa Z, Zahran MAH, Abdel Azzem M. Recent advances in silver nanoparticle-based electrochemical sensors for determining organic pollutants in water: A review. *Mater Adv* 2021;2:7350–65. <https://doi.org/10.1039/d1ma00769f>.
- [13] Zhang XF, Liu ZG, Shen W, Gurunathan S. Silver nanoparticles: Synthesis, characterization, properties, applications, and therapeutic approaches. *Int J Mol Sci* 2016;17. <https://doi.org/10.3390/ijms17091534>.
- [14] Caro C, M P, Klippstein R, Pozo D, P A. Silver Nanoparticles: Sensing and Imaging Applications. *Silver Nanopart* 2010. <https://doi.org/10.5772/8513>.
- [15] Bonsak J, Mayandi J, Thøgersen A, Stensrud Marstein E, Mahalingam U. Chemical synthesis of silver nanoparticles for solar cell applications. *Phys Status Solidi Curr Top Solid State Phys* 2011;8:924–7. <https://doi.org/10.1002/pssc.201000275>.
- [16] Blossi M, Albonetti S, Ortelli S, Costa AL, Ortolani L, Dondi M. Green and easily scalable microwave synthesis of noble metal nanosols (Au, Ag, Cu, Pd) usable as catalysts. *N J Chem* 2014;38:1401–9. <https://doi.org/10.1039/C3NJ00894K>.
- [17] Bolla PA, Huggias S, Serradell MA, Ruggera JF, Casella ML. Synthesis and catalytic application of silver nanoparticles supported on lactobacillus Kefiri S-Layer proteins. *Nanomaterials* 2020;10:1–16. <https://doi.org/10.3390/nano10112322>.
- [18] Dong XY, Gao ZW, Yang KF, Zhang WQ, Xu LW. Nanosilver as a new generation of silver catalysts in organic transformations for efficient synthesis of fine chemicals. *Catal Sci Technol* 2015;5:2554–74. <https://doi.org/10.1039/c5cy00285k>.
- [19] Yaqoob AA, Umar K, Ibrahim MNM. Silver nanoparticles: various methods of synthesis, size affecting factors and their potential applications—a review. *Appl Nanosci* 2020;10:1369–78. <https://doi.org/10.1007/s13204-020-01318-w>.
- [20] Ferrous Z, Nemmar A. Health impact of silver nanoparticles: A review of the biodistribution and toxicity following various routes of exposure 2020;vol. 21. <https://doi.org/10.3390/ijms21072375>.
- [21] Prabhu S., Poulouse E.K. Silver nanoparticles: mechanism of antimicrobial action, synthesis, medical applications, and toxicity effects 2012:1–10.
- [22] Firdhouse MJ, Lalitha P, Nanoparticles S. *Biosynthesis of Silver Nanoparticles and Its Applications* 2015;2015.
- [23] Quintero-Quiroz C, Acevedo N, Zapata-Giraldo J, Botero LE, Quintero J, Zárate-Triviño D, et al. Optimization of silver nanoparticle synthesis by chemical reduction and evaluation of its antimicrobial and toxic activity. *Biomater Res* 2019;23:1–15. <https://doi.org/10.1186/s40824-019-0173-y>.
- [24] Shahjahan M. Synthesis and Characterization of Silver Nanoparticles by Sol-Gel Technique. *Nanosci Nanometrology* 2017;3:34. <https://doi.org/10.11648/j.nsnm.20170301.16>.
- [25] Rautela A, Rani J, Debnath (Das) M. Green synthesis of silver nanoparticles from *Tectona grandis* seeds extract: characterization and mechanism of antimicrobial action on different microorganisms. *J Anal Sci Technol* 2019;10. <https://doi.org/10.1186/s40543-018-0163-z>.
- [26] Makowski T, Kowalczyk D, Fortuniak W, Brzezinski S, Kregiel D. Electrochemical deposition of silver nanoparticle and polymerization of pyrrole on fabrics via conducting multiwall carbon nanotubes. *Cellulose* 2015;22:3063–75. <https://doi.org/10.1007/s10570-015-0725-9>.
- [27] Gakiya-Teruya M, Palomino-Marcelo L, Rodriguez-Reyes JCF. Synthesis of highly concentrated suspensions of silver nanoparticles by two versions of the chemical reduction method. *Methods Protoc* 2019;2:1–5. <https://doi.org/10.3390/mps2010003>.
- [28] Gontijo LAP, Raphael E, Ferrari DPS, Ferrari JL, Lyon JP, Schiavon MA. pH effect on the synthesis of different size silver nanoparticles evaluated by dls and their size-dependent antimicrobial activity. *Rev Mater* 2020;25:1–10. <https://doi.org/10.1590/S1517-707620200004.1145>.
- [29] Jiang XC, Chen WM, Chen CY, Xiong SX, Yu AB. Role of Temperature in the Growth of Silver Nanoparticles Through a Synergetic Reduction Approach. *Nanoscale Res Lett* 2011;6:1–9. <https://doi.org/10.1007/s11671-010-9780-1>.

- [30] Amir D, Nasaruddin RR, Engliman NS, Sulaiman S, Mastuli MS. Effect of Stabilizers in the Synthesis of Silver Nanoparticles and Methylene Blue Oxidation. *IOP Conf Ser Mater Sci Eng* 2021;1192:012031. <https://doi.org/10.1088/1757-899x/1192/1/012031>.
- [31] Gentry S, Kendra S, Bezpalko M. Ostwald Ripening in Metallic Nanoparticles: Stochastic Kinetics. *J Phys Chem C* 2011;115. <https://doi.org/10.1021/jp2009786>.
- [32] Balavandy SK, Shameli K, Biak DRBA, Abidin ZZ. Stirring time effect of silver nanoparticles prepared in glutathione mediated by green method. *Chem Cent J* 2014;8:1–10. <https://doi.org/10.1186/1752-153X-8-11>.
- [33] Deshmukh AR, Gupta A, Kim BS. Ultrasound Assisted Green Synthesis of Silver and Iron Oxide Nanoparticles Using Fenugreek Seed Extract and Their Enhanced Antibacterial and Antioxidant Activities. *Biomed Res Int* 2019;2019. <https://doi.org/10.1155/2019/1714358>.
- [34] Kthiri A, Hamimed S, Othmani A, Landoulsi A, O'Sullivan S, Sheehan D. Novel static magnetic field effects on green chemistry biosynthesis of silver nanoparticles in *Saccharomyces cerevisiae*. *Sci Rep* 2021;11:1–9. <https://doi.org/10.1038/s41598-021-99487-3>.
- [35] Hasan KMF, Xiaoyi L, Shaoqin Z, Horváth PG, Bak M, Bejő L, et al. Functional silver nanoparticles synthesis from sustainable point of view: 2000 to 2023 - A review on game changing materials. *Heliyon* 2022;8. <https://doi.org/10.1016/j.heliyon.2022.e12322>.
- [36] Tolaymat TM, El Badawy AM, Genaidy A, Scheckel KG, Luxton TP, Suidan M. An evidence-based environmental perspective of manufactured silver nanoparticle in syntheses and applications: A systematic review and critical appraisal of peer-reviewed scientific papers. *Sci Total Environ* 2010;408:999–1006. <https://doi.org/10.1016/j.scitotenv.2009.11.003>.
- [37] Jeong Y, Lim DW, Choi J. Assessment of size-dependent antimicrobial and cytotoxic properties of silver nanoparticles. *Adv Mater Sci Eng* 2014;2014. <https://doi.org/10.1155/2014/763807>.
- [38] Lu Z, Rong K, Li J, Yang H, Chen R. Size-dependent antibacterial activities of silver nanoparticles against oral anaerobic pathogenic bacteria. *J Mater Sci Mater Med* 2013;24:1465–71. <https://doi.org/10.1007/s10856-013-4894-5>.
- [39] Bakand S, Hayes A. Toxicological considerations, toxicity assessment, and risk management of inhaled nanoparticles. *Int J Mol Sci* 2016;17:1–17. <https://doi.org/10.3390/ijms17060929>.
- [40] Motta G, Gualtieri M, Saibene M, Bengalli R, Briagliadori A, Carrière M, et al. Preliminary Toxicological Analysis in a Safe-by-Design and Adverse Outcome Pathway-Driven Approach on Different Silver Nanoparticles: Assessment of Acute Responses in A549 Cells. *Toxics* 2023;11. <https://doi.org/10.3390/toxics11020195>.
- [41] Karlsson HL, Cronholm P, Gustafsson J, Moeller L. Copper oxide nanoparticles are highly toxic: A comparison between metal oxide nanoparticles and carbon nanotubes. *Chem Res Toxicol* 2008;21:1726–32. <https://doi.org/10.1021/tx800064j>.
- [42] Nel A, Xia T, Mädler L, Li N. Toxic Potential of Materials at the Nanolevel. *LP – 627 Science (80-)* 2006;311:622. <https://doi.org/10.1126/science.1114397>.
- [43] Stone V, Johnston H, Clift MJD. Air pollution, ultrafine and nanoparticle toxicology: Cellular and molecular interactions. *IEEE Trans Nanobioscience* 2007;6:331–40. <https://doi.org/10.1109/TNB.2007.909005>.
- [44] Qing Y, Cheng L, Li R, Liu G, Zhang Y, Tang X, et al. Potential antibacterial mechanism of silver nanoparticles and the optimization of orthopedic implants by advanced modification technologies. *Int J Nanomed* 2018;13:3311–27. <https://doi.org/10.2147/IJN.S165125>.
- [45] Stacey N, Winder C. *Toxicology of gases, vapours and particulates. Occupational Toxicology*. 2nd ed., Boca Raton, FL, USA: CRC Press; 2004.
- [46] Hussain S.M. In vitro toxicity of nanoparticles in BRL 3A rat liver cells 2005;19: 975–983. <https://doi.org/10.1016/j.tiv.2005.06.034>.
- [47] Gojova A., Guo B., Kota R.S., Rutledge J.C., Kennedy I.M., Barakat A.I. Induction of Inflammation in Vascular Endothelial Cells by Metal Oxide Nanoparticles: Effect of Particle Composition 2007;115:403–409. <https://doi.org/10.1289/ehp.8497>.
- [48] Chithrani D.B. Intracellular uptake, transport, and processing of gold nanostructures 2010;7688. <https://doi.org/10.3109/09687688.2010.507787>.
- [49] Article O. Increased inflammation and altered macrophage chemotactic responses caused by two ultrafine particle types 2004;442–447. <https://doi.org/10.1136/oem.2003.008227>.
- [50] Kim I, Joachim E, Choi H, Kevin KK. Toxicity of silica nanoparticles depends on size, dose, and cell type. *Nanomed Nanotechnol, Biol Med* 2015;11:1407–16. <https://doi.org/10.1016/j.nano.2015.03.004>.
- [51] Poland C.A., Duffin R., Kinloch I.A.N., Maynard A., Wallace W.A.H., Seaton A., et al. Carbon nanotubes introduced into the abdominal cavity of mice show asbestos-like pathogenicity in a pilot study 2008;423–428. <https://doi.org/10.1038/nnano.2008.111>.
- [52] Bakand S., Winder C., Khalil C., Hayes A. A novel in vitro exposure technique for toxicity testing of selected volatile organic compounds w 2006:100–105. <https://doi.org/10.1039/b509812b>.
- [53] Costa A.L., Blossi M. WO2016125070A1, 2016.
- [54] Costa AL, Blossi M, Briagliadori A, Zanoni I, Ortelli S, Simeone FC, et al. Eco design for Ag-based solutions against SARS-CoV-2 and E. coli. *Environ Sci Nano* 2022;9: 4295–304. <https://doi.org/10.1039/d2en00178k>.
- [55] Thit A, Selck H, Bjerregaard HF. Toxicology in Vitro Toxic mechanisms of copper oxide nanoparticles in epithelial kidney cells. *Toxicol Vitr* 2015;29:1053–9. <https://doi.org/10.1016/j.tiv.2015.03.020>.
- [56] Warheit DB. Nanoparticles: health impacts? *Mater Today* 2004;32–5.
- [57] Sanganerla P., Sachar S., Chandra S., Bahadur D., Ray P. Cellular internalization and detailed toxicity analysis of protein-immobilized iron oxide nanoparticles 2014;2011:125–134. <https://doi.org/10.1002/jbm.b.33178>.
- [58] Mcshan D, Ray PC, Yu H. ScienceDirect Molecular toxicity mechanism of nanosilver. *J Food Drug Anal* 2014;22:116–27. <https://doi.org/10.1016/j.jfda.2014.01.010>.
- [59] Boland S., Hussain S., Baeza-squiban A. Carbon black and titanium dioxide nanoparticles induce distinct 2014;6:641–652. <https://doi.org/10.1002/wnan.1302>.
- [60] Mustajbegovic J, Zuskin E, Schachter EN. Respiratory Findings in Chemical Workers Exposed to Low Concentrations of Organic and Inorganic Air Pollutants; 2000;440:431–440.
- [61] Seaton A. Nanotechnology and the occupational physician; 2006:312–316. <https://doi.org/10.1093/occmed/kql053>.
- [62] Brunner T.J., Wick P., Bruinink A. In Vitro Cytotoxicity of Oxide Nanoparticles: Comparison to Asbestos, Silica, and the Effect of Particle Solubility †; 2006;40: 4374–4381.
- [63] Khandoga A, Stoeger T, Khandoga AG, Bihari P, Karg E, Ettehadieh D, et al. Platelet adhesion and fibrinogen deposition in murine microvessels upon inhalation of nanosized carbon particles. *J Thromb Haemost* 2010;8:1632–40. <https://doi.org/10.1111/j.1538-7836.2010.03904.x>.
- [64] Oliveira JCA, Frey J, Zhang S-Q, Xu L-C, Li X, Li S-W, et al. When machine learning meets molecular synthesis. *Trends Chem* 2022;4:863–85. <https://doi.org/10.1016/j.trechm.2022.07.005>.
- [65] Cova TFGG, Pais AAC. Deep learning for deep chemistry. *Optim Predict Chem Patterns* 2019;7:1–22. <https://doi.org/10.3389/chem.2019.00809>.
- [66] Reza G. A predictive model on size of silver nanoparticles prepared by green synthesis method using hybrid artificial neural network-particle swarm optimization algorithm. *Measurement* 2020;151:107199. <https://doi.org/10.1016/j.measurement.2019.107199>.
- [67] Pellegrino F, Isopescu R, Pellutiè L, Sordello F, Rossi AM, Ortel E, et al. Machine learning approach for elucidating and predicting the role of synthesis parameters on the shape and size of - nanoparticles. *Sci Rep* 2020;1–11. <https://doi.org/10.1038/s41598-020-75967-w>.
- [68] Winkler DA. Role of artificial intelligence and machine learning in nanosafety. *Small* 2020;16. <https://doi.org/10.1002/sml.202001883>.
- [69] Mekki-Berrada F, Ren Z, Huang T, Wong WK, Zheng F, Xie J, et al. Two-step machine learning enables optimized nanoparticle synthesis. *Npj Comput Mater* 2021;7:1–10. <https://doi.org/10.1038/s41524-021-00520-w>.
- [70] Gain U, Hottl V. Low-code AutoML-augmented Data Pipeline – a review and experiments. *J Phys Conf Ser* 2021;1828:12015. <https://doi.org/10.1088/1742-6596/1828/1/012015>.
- [71] Mangalathu S, Hwang S-H, Jeon J-S. Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Eng Struct* 2020;219:110927. <https://doi.org/10.1016/j.engstruct.2020.110927>.
- [72] Anjum M, Khan K, Ahmad W, Ahmad A, Amin MN, Nafees A. New SHapley additive ExPlanations (SHAP) approach to evaluate the raw materials interactions of steel-fiber-reinforced concrete. *Materials* 2022;15. <https://doi.org/10.3390/ma15186261>.
- [73] Badiah HI, Seede F, Supriyanto G, Zaidan AH. Synthesis of silver nanoparticles and the development in analysis method. *IOP Conf Ser Earth Environ Sci* 2019; 217. <https://doi.org/10.1088/1755-1315/217/1/012005>.
- [74] Sukweenadhi J, Setiawan KI, Avanti C, Kartini K, Rupa EJ, Yang DC. Scale-up of green synthesis and characterization of silver nanoparticles using ethanol extract of *Platago major* L. leaf and its antibacterial potential. *South Afr J Chem Eng* 2021;38:1–8. <https://doi.org/10.1016/j.sajce.2021.06.008>.
- [75] Velgosova O, Mačák L, Čizmarová E, Mára V. Influence of reagents on the synthesis process and shape of silver nanoparticles. *Materials* 2022;15. <https://doi.org/10.3390/ma15196829>.
- [76] Galatage ST, Hebalkar AS, Dhobale SV, Mali OR, Kumbhar PS, Nikade SV, et al. In: Kumar S, Kumar P, Pathak CS, editors. *Silver Nanoparticles: Properties, Synthesis, Characterization, Applications and Future Trends*. Rijeka: IntechOpen; 2021. <https://doi.org/10.5772/intechopen.99173>.
- [77] Yin IX, Zhang J, Zhao IS, Mei ML, Li Q, Chu CH. The antibacterial mechanism of silver nanoparticles and its application in dentistry. *Int J Nanomed* 2020;15: 2555–62. <https://doi.org/10.2147/IJN.S246764>.
- [78] Anees Ahmad S, Sachi Das S, Khatoun A, Tahir Ansari M, Afzal M, Saquib Hasnain M, et al. Bactericidal activity of silver nanoparticles: a mechanistic review. *Mater Sci Energy Technol* 2020;3:756–69. <https://doi.org/10.1016/j.mset.2020.09.002>.
- [79] Vi TTT, Kumar SR, Huang YT, Chen DW, Liu YK, Lue SJ. Size-dependent antibacterial activity of silver nanoparticle-loaded graphene oxide nanosheets. *Nanomaterials* 2020;10:1–18. <https://doi.org/10.3390/nano10061207>.
- [80] Loza K, Epple M. Silver nanoparticles in complex media: an easy procedure to discriminate between metallic silver nanoparticles, reprecipitated silver chloride, and dissolved silver species. *RSC Adv* 2018;8:24386–91. <https://doi.org/10.1039/c8ra04500c>.
- [81] Loza K, Diendorf J, Sengstock C, Ruiz-Gonzalez L, Gonzalez-Calbet JM, Vallet-Regi M, et al. The dissolution and biological effects of silver nanoparticles in biological media. *J Mater Chem B* 2014;2:1634–43. <https://doi.org/10.1039/c3tb21569e>.
- [82] Lin Y-S, Tsai C-P, Huang H-Y, Kuo C-T, Hung Y, Huang D-M, et al. Well-ordered mesoporous silica nanoparticles as cell markers. *Chem Mater* 2005;17:4570–3. <https://doi.org/10.1021/cm051014c>.

- [83] Islam MA, Jacob MV, Antunes E. A critical review on silver nanoparticles: from synthesis and applications to its mitigation through low-cost adsorption by biochar. *J Environ Manag* 2021;281:111918. <https://doi.org/10.1016/j.jenvman.2020.111918>.
- [84] Siddiqi KS, Husen A, Rao RAK. A review on biosynthesis of silver nanoparticles and their biocidal properties. *J Nanobiotechnol*. 2018;16. <https://doi.org/10.1186/s12951-018-0334-5>.
- [85] Zanon I, Crosera M, Pavoni E, Adami G, Mauro M, Costa AL, et al. Use of single particle ICP-MS to estimate silver nanoparticle penetration through baby porcine mucosa. *Nanotoxicology* 2021;15:1005–15. <https://doi.org/10.1080/17435390.2021.1940338>.
- [86] Mirzaei M, Furxhi I, Murphy F, Mullins M. A machine learning tool to predict the antibacterial capacity of nanoparticles. *Nanomaterials* 2021;11. <https://doi.org/10.3390/nano11071774>.
- [87] Mirzaei M, Furxhi I, Murphy F, Mullins M. A supervised machine-learning prediction of textile's antimicrobial capacity coated with nanomaterials. *Coatings* 2021;11:1–18. <https://doi.org/10.3390/coatings11121532>.
- [88] Mohammed AE, Alghamdi SS, Alotaibi MO, Rahman I, Shami A, Suliman RS, et al. In silico prediction of malva viscus arboreus metabolites and green synthesis of silver nanoparticles – opportunities for safer anti-bacterial and anti-cancer precision medicine. *Int J Nanomed* 2023;18:2141–62. <https://doi.org/10.2147/IJN.S400195>.
- [89] Botte E, Vagaggini P, Costa J, Faccani L, Zanon I, Costa A, et al. A novel in silico tool for dose assessment in cell culture nanotoxicology. *Biomed Sci Eng* 2021;4 (s1):39–40. <https://doi.org/10.4081/bse.2021>.
- [90] Furxhi I, Bengalli R, Motta G, Mantecca P, Kose O, Carriere M, et al. Data-driven quantitative intrinsic hazard criteria for nanoproduct development in a safe-by-design paradigm: a case study of silver nanoforms. *ACS Appl Nano Mater* 2023;6: 3948–62. <https://doi.org/10.1021/acsanm.3c00173>.
- [91] Mao BH, Luo YK, Wang BJ, Chen CW, Cheng FY, Lee YH, et al. Use of an in silico knowledge discovery approach to determine mechanistic studies of silver nanoparticles-induced toxicity from in vitro to in vivo. *Part Fibre Toxicol* 2022; 19:1–25. <https://doi.org/10.1186/s12989-022-00447-0>.
- [92] Furxhi I, Varesano A, Salman H, Mirzaei M, Battistello V, Tomasoni IT, et al. Data shepherding in nanotechnology: an antimicrobial functionality data capture template. *Coatings* 2021;11:1–18. <https://doi.org/10.3390/coatings11121486>.
- [93] Stavinskaya O, Laguta I, Fesenko T, Krumova M. EFFECT of temperature on green synthesis of silver nanoparticles using Vitex agnus-castus extract. *Chem J Mold* 2019;14:117–21. <https://doi.org/10.19261/cjm.2019.636>.
- [94] Xiliang Q, Yang C, Tiesong L, Peng H, Jun W, Ping L, et al. Large-scale synthesis of silver nanoparticles by aqueous reduction for low-temperature sintering bonding. *J Nanomater* 2014;2014:594873. <https://doi.org/10.1155/2014/594873>.
- [95] Fu LM, Hsu JH, Shih MK, Hsieh CW, Ju WJ, Chen YW, et al. Process optimization of silver nanoparticle synthesis and its application in mercury detection. *Micromachines* 2021;12:1–16. <https://doi.org/10.3390/mi12091123>.
- [96] Furxhi I. Health and environmental safety of nanomaterials: O Data, Where Art Thou? *NanoImpact* 2022;25:100378. <https://doi.org/10.1016/j.impact.2021.100378>.
- [97] Furxhi I, Murphy F, Mullins M, Arvanitis A, Poland CA. Nanotoxicology data for in silico tools: a literature review. *Nanotoxicology* 2020;14:612–37. <https://doi.org/10.1080/17435390.2020.1729439>.
- [98] Furxhi I, Murphy F, Mullins M, Arvanitis A, Poland CA. Practices and trends of machine learning application in nanotoxicology. *Nanomaterials* 2020;10:1–32. <https://doi.org/10.3390/nano10010116>.
- [99] Furxhi I, Willighagen E, Evelo C, Costa A, Gardini D, Ammar A. A data reusability assessment in the nanosafety domain based on the NSDRA framework followed by an exploratory quantitative structure activity relationships (QSAR) modeling targeting cellular viability. *NanoImpact* 2023;31:100475. <https://doi.org/10.1016/j.impact.2023.100475>.
- [100] Lebre F, Chatterjee N, Costa S, Fernández-De-gortari E, Lopes C, Meneses J, et al. Nanosafety: an evolving concept to bring the safest possible nanomaterials to society and environment. *Nanomaterials* 2022;12. <https://doi.org/10.3390/nano12111810>.
- [101] Dumit VI, Ammar A, Bakker MI, Bañares MA, Bossa C, Costa A, et al. From principles to reality. FAIR implementation in the nanosafety community. *Nano Today* 2023;51:101923. <https://doi.org/10.1016/j.nantod.2023.101923>.
- [102] Furxhi I, Perucca M, Blosi M, Lopez de Ipiña J, Oliveira J, Murphy F, et al. ASINA project: towards a methodological data-driven sustainable and safe-by-design approach for the development of nanomaterials. *Front Bioeng Biotechnol* 2022;9: 1–8. <https://doi.org/10.3389/fbioe.2021.805096>.
- [103] Scheffler M, Aeschlimann M, Albrecht M, Bereau T, Bungartz HJ, Felser C, et al. FAIR data enabling new horizons for materials research. *Nature* 2022;604: 635–42. <https://doi.org/10.1038/s41586-022-04501-x>.