

## GestaltMatcher research study and GestaltMatcher database.

### Study information

Dear Sir or Madam,

Thank you for your interest in the GestaltMatcher Research Study and Database (GMDB)! Your participation in this study is voluntary. Before providing data/photos/further information, please make sure you have read the following study information and have given the necessary consent.

### Study Information

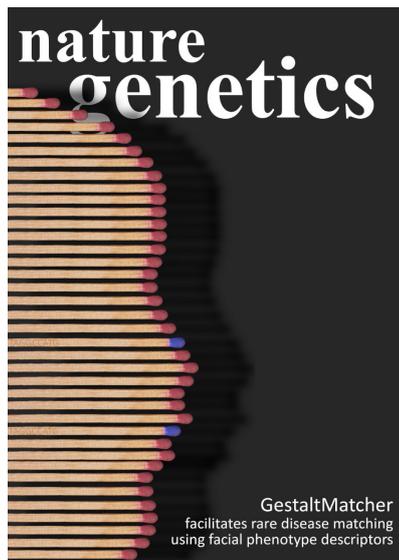
You or your child has or is suspected of having a rare genetic disorder. The diagnostic process of such rare genetic disorders can be lengthy. Often, affected individuals and their families experience a diagnostic odyssey. Our research study aims to shorten this diagnostic process.

In many genetic disorders, facial features that are specific for the disorder provide a hint of the disease. A good example for this is Down syndrome. Many people immediately recognize an affected person by his or her face, because patients share some special characteristics. A specialized doctor can recognize many more and rarer diseases. However, a prerequisite for this is an appropriate and sufficient experience with the individual clinical patterns. But it is precisely this experience that is often lacking in the case of rare diseases.



Figure 1: Facial similarities shown in the example of Down's syndrome.

Modern computer-aided image analysis methods can also learn to recognize such features in the face and thus reliably calculate similarities to individuals with an already confirmed diagnosis. In the PEDIA study, published in the renowned journal *Genetics in Medicine* in 2019,



we were already able to show that the diagnosis rate is significantly increased by combining it with image analysis. Like a doctor, however, the computer system must first learn the characteristics of the diseases in order to recognize them. With the GestaltMatcher algorithm, published in 2022 in the high-impact journal *Nature Genetics*, we have developed an artificial intelligence (AI) that can not only recognize diseases it has already learned, but also diseases it is not previously trained to recognize. It also requires significantly fewer patient images than comparable systems to make the correct diagnosis. Such computational methods can provide medical professionals with a clue to a genetic disorder and assist in selecting a particular molecular test. In addition, AI can also be used to interpret the results of molecular genetic testing.

Our research study will investigate and improve the quality of currently available image analysis methods (e.g., GestaltMatcher). We will also investigate whether a new mutation assessment protocol called PEDIA (prioritization of exome data by image analysis) improves diagnostic yield.

The performance of the GestaltMatcher AI depends substantially on the dataset it was trained with. For this purpose, we have launched the GestaltMatcher database, where the gathered data is collected and made searchable. It is also possible to use the GestaltMatcher AI to determine similarities between individuals in the database. This database is and will continue to be accessible exclusively to a selected scientific audience for the purpose of improving the diagnosis of rare diseases. The carrier of this database is the non-profit organization Arbeitsgemeinschaft Gen-Diagnostik (AGD) e.V. A scientific analysis of the data is performed at the Institute of Genomic Statistics and Bioinformatics (IGSB) at the University of Bonn, Germany.

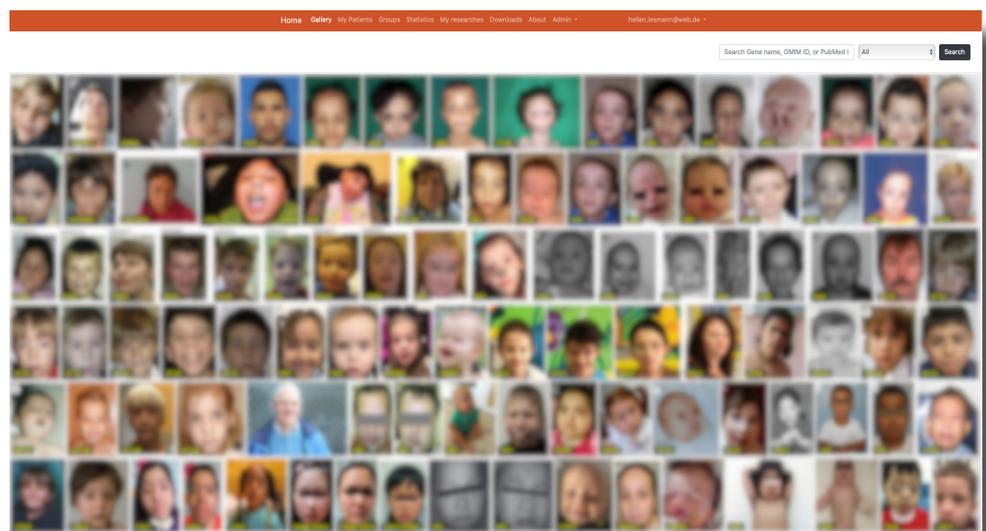


Figure 3: Gallery view

Doctors can use the gallery to search for images of specific diseases or disease-causing genes. This way, they can see many portrait images of a syndrome at a glance, compare them with their patients and gain experience. In the future, they may then be able to recognize rare diseases more quickly on the basis of particular facial features.

## Potential risks

As part of routine diagnostics, the clinical features of patients are documented using photographs and medical genetic terminology. These characteristics, the medical photographs and the test results from the laboratory are transmitted to us by you or your physician and stored in the GestaltMatcher database in pseudonymized form. Pseudonymization means that other users of the platform do not know the name, address or other personal data of the study participant. The clinical data can only be assigned to a person via the pseudonym, which is known only to you and the submitting doctor.

However, there are confidentiality risks associated with any collection, storage, use, and transmission of data (e.g., the possibility of identifying the individual). All data of the research study are stored on a server at the IGSB in Bonn, which complies with current data protection regulations. Nevertheless, these risks cannot be completely excluded.

Since our study only collects data from routine diagnostics, there are no medical risks associated with participation.

## Potential benefits from your participation in the study

Our study aims to improve the diagnosis of genetic diseases and to find the correct diagnosis more quickly. In the future, this may prevent unnecessary testing and help patients receive appropriate medical care for their specific disease sooner. In the long run, this may also contribute to their successful treatment. If you do not have a diagnosis yet, participation in this study could potentially shorten your diagnostic process. However, there is no direct therapeutic benefit to you or your child from participating in this project. If no possible diagnosis is found, exchange with other international platforms similar to GMDB is also possible after your approval. If the diagnosis is known prior to participation, you may help improve care for other patients who are affected by the same condition.

The screenshot displays the GestaltMatcher database interface for a specific patient. It is organized into several sections:

- Patient Information:** Case ID: 4225, Clinicians Reference: HL\_533, User: Ms. Hellen Lesmann, Ethnicity: Asian - South/Indian, Ethnicity note, Gender: female, Group: NONE, Note: Microdeletion.
- Ethical, Legal, and Social Aspects (ELSA):** PubMed, DOI, Consent obtained, Family numbering, Subject numbering, Corresponding author or clinician that obtained informed consent, Email, and an "Add another patient" button.
- Photos:** An "Upload Photo" button and a table of existing photos.
- Diagnosed disorders:** A table listing disorders with OMIM, Disorder name, and Diagnosed status.
- Molecular Information:** A table listing genes, tests, and HGVS details.
- Phenotypic Information:** A section indicating "No phenotypic features!".

Photo	File name	Type	Age	Age note	Which person	Score	Updated date	
	Bildschirmfoto 2021-06-05 um 17.59.45.png	Frontal face	3.0	-	Index	Important	2021-06-05	

OMIM	Disorder	Diagnosed
182290	SMITH-MAGENIS SYNDROME	Molecularly diagnosed

Gene	Test	HGVS
RAI1 10743	Microarray None	

Figure 4: Representation of individual participants in the database.

No personal data is collected in the database. You can see how the data is represented in this figure. The following are stored: gender, ethnicity, the disease, the causative disease gene, and other phenotype information, if applicable. The uploaded photo will be displayed in the „Photos“ section and in the gallery.

We intend to publish the results of this research project in scientific journals and at conferences. These publications will not contain any personal information that would identify you. Participation in the study is free of charge, and you are not entitled to any compensation, other financial benefits, or profits that may be obtained based on this research.

## Can I withdraw my consent?

You can withdraw your consent at any time in writing or verbally without giving any reason and without any disadvantage to you. If you revoke your consent, you can also request the deletion of your data.

## Do you have questions, feedback or suggestions?

Please contact us: [info@gestaltmatcher.org](mailto:info@gestaltmatcher.org)



**Peter Krawitz, MD/PhD**

Prof. Dr. med. Peter Krawitz

Head at the Institute for Genomic Statistics and Bioinformatics  
Chairman of the Arbeitsgemeinschaft für Gen-Diagnostik e.V. (AGD)  
University Hospital Bonn  
Rheinische Friedrich-Wilhelms-University Bonn  
Venusberg Campus 1  
53127 Bonn, Germany

web: <http://www.igsb.uni-bonn.de>



**Tzung-Chien Hsieh**



**Tom Kamphans, PhD**



**Hellen Lesmann, MD**



**Tori Pantel, MD**



**Shahida Moosa, MD/PhD**



**Wolfgang Meiswinkel**



**Stanislav Rosnev**



**Jing-Mei Li, PhD**