The magazine of the
Center for Molecular Fingerprinting

WAVES

PATH TO A HEALTHIER FUTURE

2023

Impressum

Published by:

Center for Molecular Fingerprinting Research Nonprofit LCC.

Prof. Dr. Ferenc Krausz, CEO

H-1093 Budapest, Czuczor utca 2-10.

Designed and edited by: Lounge Design

Photo credit: Thorsten Naeser

Printed by: Crew Kft.

Table of contents

Dear Reader,

Welcome to the first issue of Waves, the magazine of the Center for Molecular Fingerprinting (CMF).

The aim of the magazine is to provide information for the broader scientific community as well as the general public about our bold vision for the future of healthcare. It is about how we contribute to making this vision a reality, and about the vast frontiers we need to cross to be successful. By reading the magazine, you will get to know our interdisciplinary team and how we are taking concerted action to achieve our goals. We will use this platform to report on the progress we have made on putting our novel concepts for health monitoring to the test in real life.

We are on a journey towards creating the prerequisites for a new kind of healthcare, one that focuses on preserving people's health. In contrast with healthcare today, where most of the resources are spent on treating diseases, many of which are diagnosed only after alarming symptoms appear, and often too late. Future preventive healthcare – which we aim to contribute to – will proactively protect your health; improving quality of life and extending the number of years people can enjoy their lives for. Proactively protecting health will save resources and help the economy and society prosper.

The future is already here with more than 10,000 health-conscious and helpful individuals participating in our Health for Hungary - Hungary for Health (H4H) Clinical Study.

Photo credit: Árpád Földházi / Mandiner

As part of the H4H Program, volunteers can monitor their health as we assess their laboratory blood count whenever they provide a blood sample. The clinical laboratory parameters we routinely analyze have already provided information on numerous conditions that are not yet symptomatic but – if they had remained undetected – might have had severe long-term consequences. Detected at an early stage, they may be prevented by lifestyle changes and appropriate medication.

However, standard clinical laboratory parameters are not enough to provide reliable early warnings for severe chronic diseases, such as cancer, metabolic and respiratory disorders and cardio-vascular diseases, which cause the majority of deaths worldwide, many of them far too early. Most of these could be efficiently fought if recognized at an early stage, the earlier the better. This is only feasible by populational health screening, using a method that provides comprehensive information and is affordable on this scale.

In a joint effort with the Ludwig-Maximilians-Universität München, CMF is developing and evaluating a new laser-based method that offers the potential to meet these requirements: infrared electric field molecular fingerprinting of blood samples. Our hope is that the amount of information provided by the new approach will, in the long run, exceed that of currently available clinical chemistry parameters accessible in standard of care. This increased information may provide sufficiently early warning for a range of conditions, long before symptoms appear. The H4H Program is designed to prove this capacity for some selected chronic diseases, specifically non-communicable diseases (NCDs) with the highest mortality, such as lung cancer, cardiovascular diseases and type 2 diabetes. If the validation succeeds, the H4H clinical network will provide all the prerequisites for turning this clinical study into the first nationwide health screening program based on infrared electric field molecular fingerprinting.

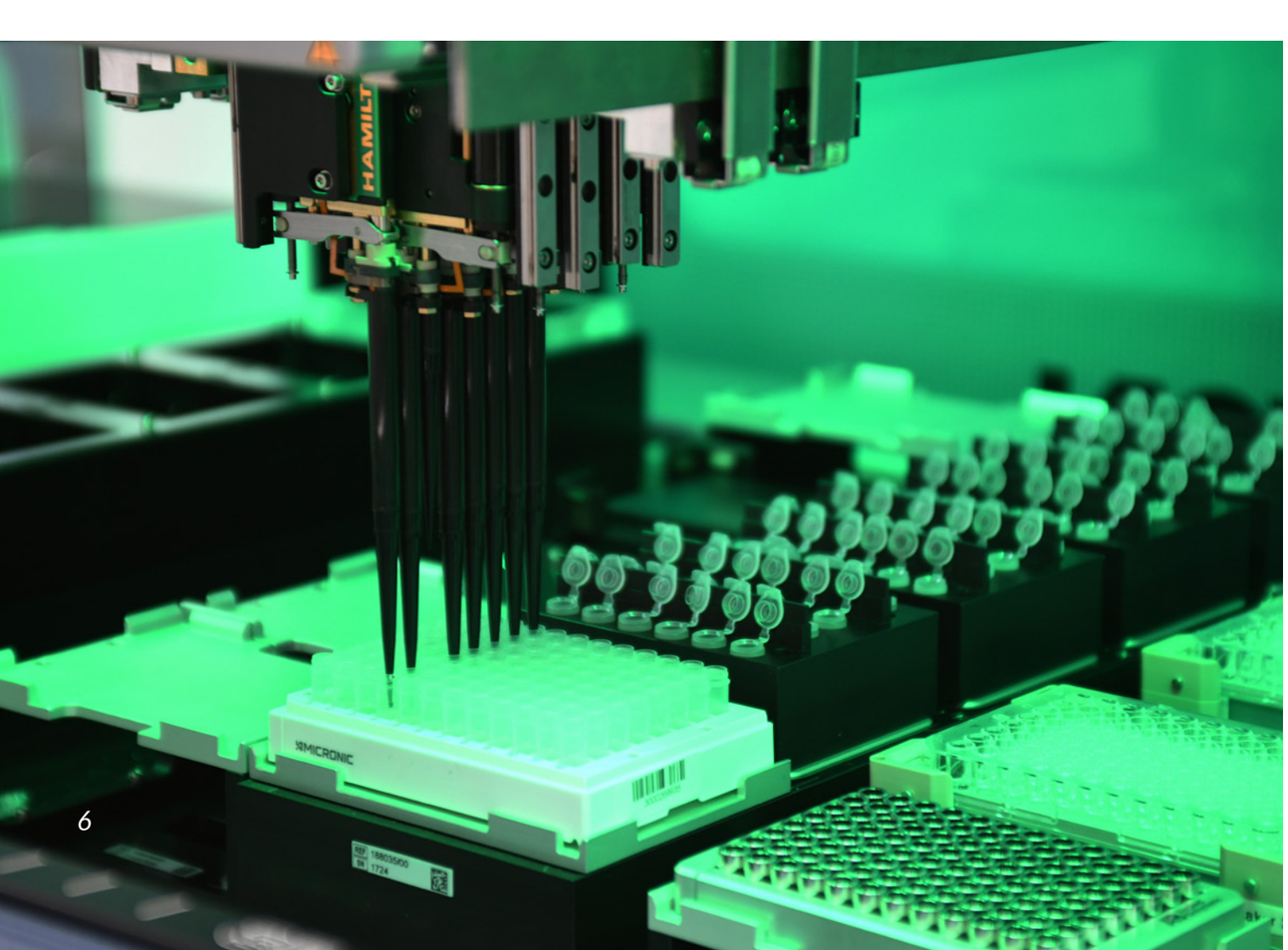
We would not be able to pursue this ambitious endeavor without the support of H4H participants, our medical partners, and, not least, the supporters of our research project. Together, we have the chance to shape healthcare for the benefit of future generations.

Sincerely yours,
Prof. Dr. Ferenc Krausz

CEO and Scientific Director,
Center for Molecular Fingerprinting
Budapest, Szeged, Munich

ABOUT US

The Center for Molecular Fingerprinting (CMF) is an interdisciplinary nonprofit research institution. We are an international team of medical doctors, nurses, laser scientists, molecular biologists, engineers, and data analysts, driven by a common goal: Moving the frontiers of monitoring and probing human health.



History

Three pioneers in ultrafast laser physics, Anne L'Huillier, Paul Corkum, and Ferenc Krausz, showed that it is possible to observe and control the motion of electrons in atoms, molecules, and solids with ultrashort pulses of light on a timescale in the range of about one hundred attoseconds.

The laser systems that L'Huillier, Corkum, and Krausz developed produce ultrashort, attosecond pulses of light that act like a camera with a shutter speed so fast that they can capture the motion of the electron in a hydrogen atom, which takes 150 attoseconds to orbit the nucleus.

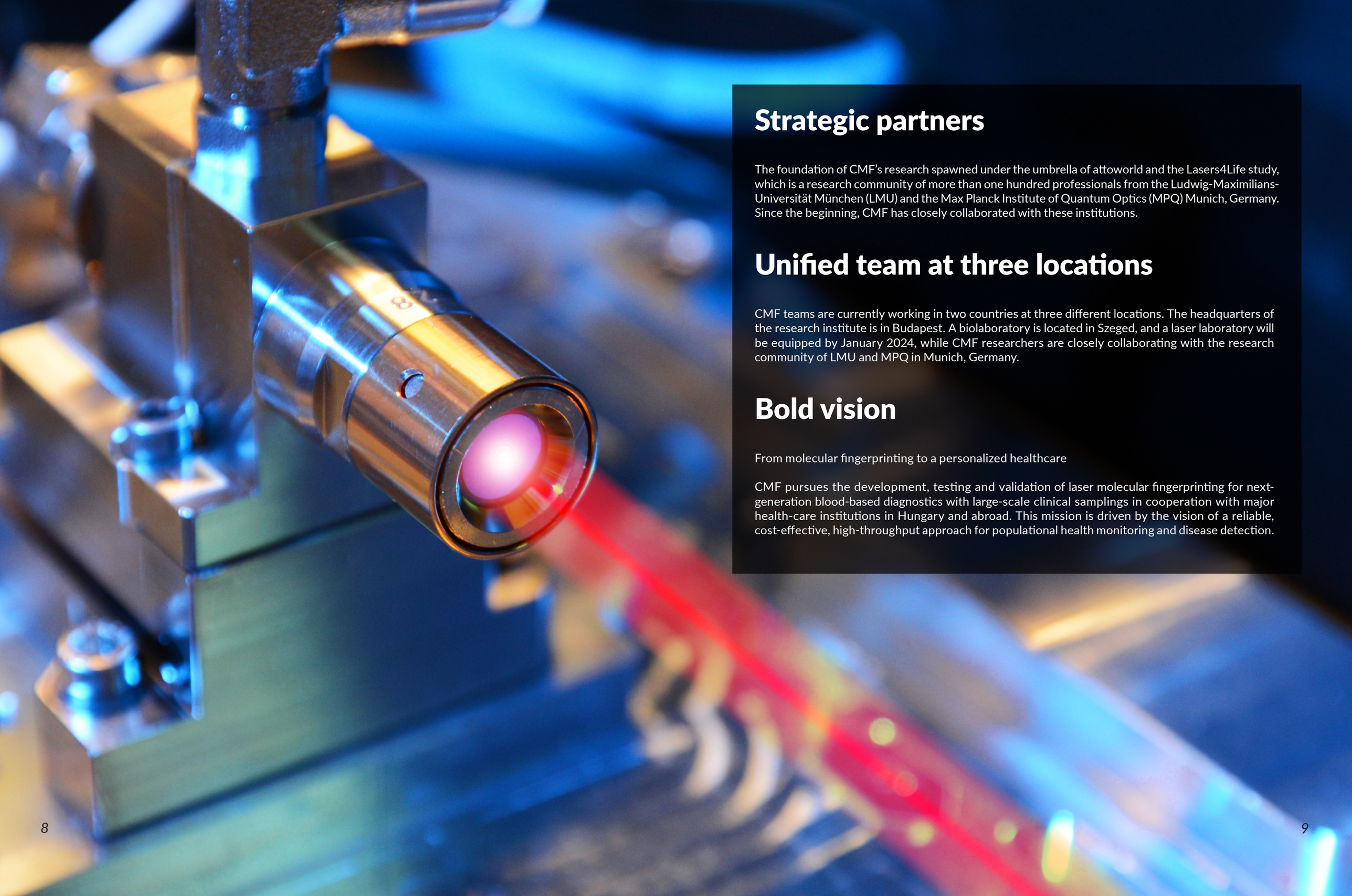
This breakthrough, based on a series of findings by distinguished researchers, was achieved in Vienna in 2001. The successful experiment not only opened the door to the detailed observation of electron motion, but was also able to confirm a number of predictions that had been made decades before by theoretical physics, but that until then had not been verifiable in the laboratory.

Once attosecond physics has revealed its undeniable potential, the plan is to use it to probe more deeply into the mysteries of matter that make up all of nature and to develop applications in fields such as electronics and biomedicine. In biomedicine, there is potential to explore the use of attosecond pulses to detect diseases. Once all the cells are removed from a blood sample, the fluid that is left is blood plasma. The molecules of blood plasma thus provide valuable clues about the health of the person the blood came from. The Center for Molecular Fingerprinting came into being to explore ways of using ultrashort pulses to obtain this information.

The research institute was established in 2019, and since then it has grown continuously. From the beginning we have operated in two locations, of which most of the scientific research has been carried out in Garching, Germany. Our flagship project, the H4H Program, has been implemented in Hungary, with the first sample being taken in July 2021. Gaining a biobank license in July 2022 was an important step forward in CMF's history. Our interim biobank at Szeged in ELI-ALPs is equipped with two -80 °C ultradeep freezers with 65,000 aliquots storage capacity each. The biolaboratory started its operations in July 2023, while the laser laboratory in Szeged will be equipped by January 2024, from which point research will be carried out in Szeged in addition to the laboratories in Garching.

Scientific mission

At the Center for Molecular Fingerprinting, innovation is driven by a vision: We define the frontiers of probing human health based on the technological, ultrashort pulsed laser development of molecular fingerprinting for the next generation of molecular diagnostics.



Strategic partners

The foundation of CMF's research spawned under the umbrella of attoworld and the Lasers4Life study, which is a research community of more than one hundred professionals from the Ludwig-Maximilians-Universität München (LMU) and the Max Planck Institute of Quantum Optics (MPQ) Munich, Germany. Since the beginning, CMF has closely collaborated with these institutions.

Unified team at three locations

CMF teams are currently working in two countries at three different locations. The headquarters of the research institute is in Budapest. A biolaboratory is located in Szeged, and a laser laboratory will be equipped by January 2024, while CMF researchers are closely collaborating with the research community of LMU and MPQ in Munich, Germany.

Bold vision

From molecular fingerprinting to a personalized healthcare

CMF pursues the development, testing and validation of laser molecular fingerprinting for next-generation blood-based diagnostics with large-scale clinical samplings in cooperation with major health-care institutions in Hungary and abroad. This mission is driven by the vision of a reliable, cost-effective, high-throughput approach for populational health monitoring and disease detection.

Introduction to the team

Prof. Dr. Ferenc Krausz

– CEO

Dr. Ferenc Krausz embarked on his journey by graduating from the Technical University Budapest in 1985. His passion for pushing the boundaries of science led him to pursue a doctorate in Laser Physics at the Technische Universität (TU) Vienna, culminating in his Ph.D. in 1991. Dr. Krausz's relentless pursuit of knowledge and innovation didn't stop there. In 1993, he habilitated in the same field. His remarkable journey continued as he assumed the role of an assistant professor in 1998, swiftly followed by a full professorship in 1999. In 2003, Dr. Krausz was appointed as the Director of the Max Planck Institute of Quantum Optics (MPQ) in Garching. His visionary leadership and groundbreaking research continued to shape the realm of laser physics. In 2004,Dr. Krausz's commitment to advancing the boundaries of laser physics led him to become a professor at the Faculty of Physics at Ludwig-Maximilians-Universität München (LMU). Since then, he has held the Chair of Experimental Physics – Laser Physics.

In a series of experiments performed between 2001 and 2004, Prof. Krausz and his team succeeded in producing and measuring attosecond light pulses and applying them for the first real-time observation of atomic-scale electron motions. These achievements earned him the reputation as the co-founder of the field of attosecond physics (along with Paul Corkum and Anne L'Huillier), a scientific discipline devoted to real-time observation and control of electron phenomena. This renown was also acknowledged by the three scientists being selected as 2015 Thomson Reuters Citation Laureates. More recently, Ferenc Krausz turned his attention to capitalizing on ultrafast laser techniques for disease detection by the molecular fingerprinting of human bio-fluids. For his contributions to ultrafast science, Prof. Dr. Ferenc Krausz shared the 2013 King-Faisal Prize for Science with Dr. Paul Corkum and the 2022 Wolf Prize in Physics with Prof. Anne L'Huillier and Dr. Paul Corkum. In June 2023, he received the "Frontiers of Knowledge Prize" donated by the Spanish Banking Corp. BBVA, also together with Prof. Anne L'Huillier and Dr. Paul Corkum.

In 2023, together with Pierre Agostini and Anne L'Huillier, he was awarded the Nobel Prize in Physics for "experimental methods that generate attosecond pulses of light for the study of electron dynamics in matter."

Dr. Mihaela Žigman

– Research Director

Dr. Mihaela Žigman is a molecular cell biologist with a strong academic background. She pursued her studies in molecular biology at the University of Ljubljana and completed her doctoral studies in molecular cell biology at the University of Vienna. It was during her post-doctoral research at the Institute of Molecular Biotechnology (IMBA) in Vienna that she fully recognized the significance of addressing fundamental biological questions through an interdisciplinary approach. With a goal of understanding how abnormalities in cell division contribute to organ growth, she embarked on a new journey by joining the Fred Hutchinson Cancer Research Center (FHCRC) in Seattle. Following her time at the University of Heidelberg, she relocated to Munich. Dr. Mihaela Žigman's utter fascination with technology-driven interdisciplinary approaches eventually led her to assume the role of a research group leader at the Ludwig-Maximilians-Universität München (LMU), as well as a research director at the Center for Molecular Fingerprinting. Her primary objective is to develop infrared molecular fingerprinting as an innovative *in vitro* medical diagnostic platform, enabling the probing of physiological phenotypes to detect the earliest transitions from a state of health to one of disease.

Laser Science

The Laser Science Division team develops technology and instrumentation to capture the infrared electric field response as it is emitted by the molecules dissolved in blood plasma. Our spectroscopy instruments take advantage of new ultrafast infrared laser technology that breaks limits in pulse duration, spectral coverage, brightness, and stability and combines them with the extreme sensitivity of field-resolved detection. To ensure high-throughput human health monitoring based on field-resolved infrared molecular fingerprinting, we employ state-of-the-art engineering approaches. We focus on achieving the stability and long-term measurement reproducibility essential for this purpose.

Life Science

The interdisciplinary team within the Life Science Division is responsible for designing, planning, and coordinating research activities related to monitoring human health. Our focus is on developing innovative methods to parameterize medically relevant information present in the non-cellular components of human blood. This work contributes to the advancement of *in vitro* disease detection and health assessment. Within the framework of the Health for Hungary – Hungary for Health (H4H) Clinical Study, we employ infrared molecular fingerprinting to analyze human blood plasma. Our primary goal is to create an effective approach for assessing human health and identifying potential deviations. This, in turn, supports medical decision-making and enhances the well-being of adult populations.

Munich

Budapest

Szeged

Data Science

The Data Science Division team explores the extent to which medical information acquired from photonic data can be utilized in medical diagnostics, personalized health monitoring, and life sciences. Specifically, we investigate relevant procedures for experimental and study design as well as data preprocessing and combine them with best practices in medical statistics and machine learning in appropriate data-science pipelines.

Biofacility

The Biofacility Division is dedicated to the creation, development, and maintenance of a standardized and homogenized biorepository for human blood plasma. Our team oversees the quality control of blood plasma, its storage within the biobank, sample management, and the coordination of sample transportation.

Clinical Research and Medical Science

The Clinical Research Division plays a crucial role in developing a strategy for implementing CMF's clinical research. This team's responsibilities include identifying potential new research sites, establishing collaboration with these institutions, and coordinating, supporting, and monitoring the execution of trials. Project coordinators are instrumental in ensuring adherence to Standard Operating Procedures (SOPs) and regulatory compliance at the research sites, with a particular focus on data integrity and validity. They also manage tasks related to regulatory licensing and license modification.

The Medical Sciences Division holds the responsibility of providing medical expertise for CMF's research activities aimed at developing an *in vitro* diagnostic tool for human health monitoring and disease detection using infrared molecular fingerprinting. The clinical and medical team contributes to the expansion of our research plans, the development of clinical study protocols, and monitoring the progress of ongoing study plans. Team members address medical inquiries and analyze data, drawing medical conclusions based on the collected information.

The finest fractions of time

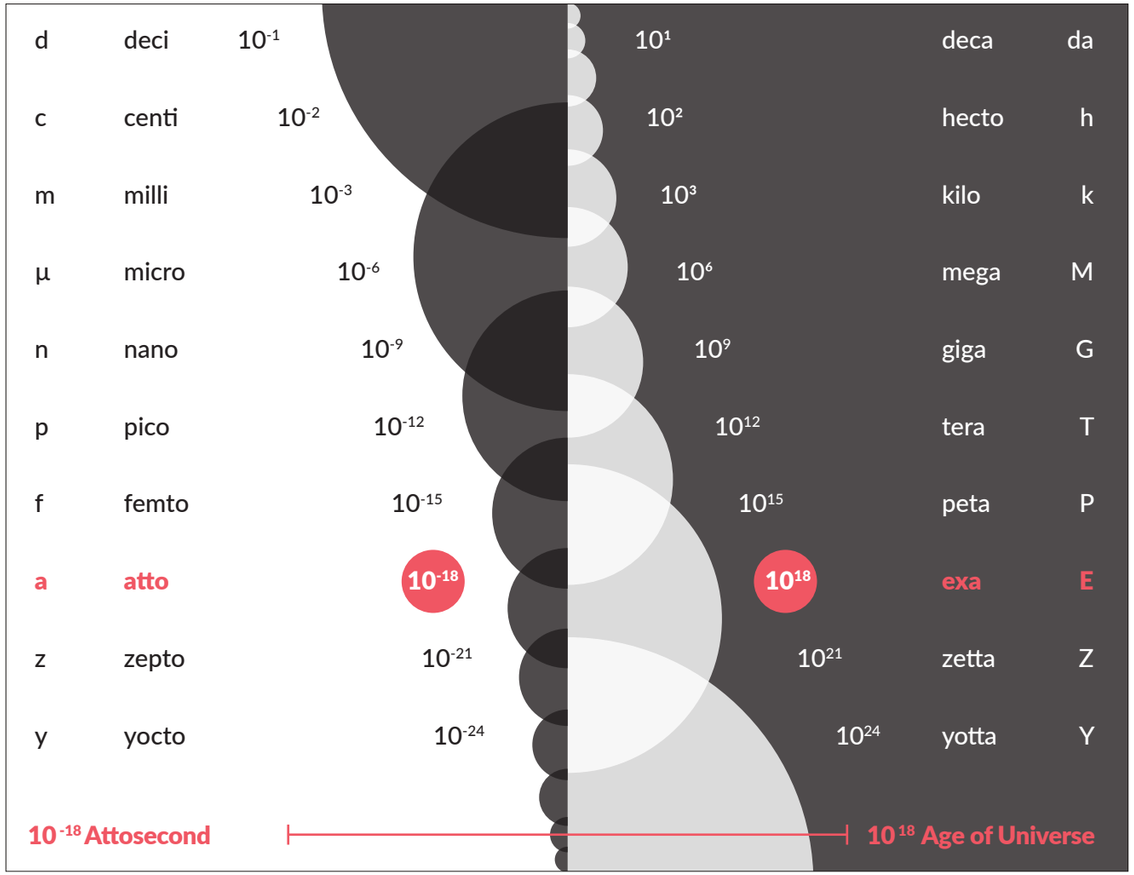
Thorsten Naeser

The quantum world challenges our everyday intuitions, including our notions of time. Electronic transitions in atoms take place within attoseconds. An attosecond is a billionth of a billionth of a second (10^{-18} sec). Prof. Dr. Ferenc Krausz and the scientists in the attoworld research group in Garching, run jointly by the Max Planck Institute of Quantum Optics (MPQ) and Ludwig-Maximilians-Universität München (LMU), study electron motions using ultrashort light pulses of this order.

In the world of electrons, things happen very, very fast. The faster objects move, the harder it becomes to observe them. A human heartbeat lasts for about a second – a timespan readily visible to the eye. To capture changes that take less than 0.2 sec, we have to use technologies that slow time down – by freezing the motions at shorter intervals – as in time-lapse photography.

This principle allows us to follow the millisecond wingbeats of flies. In order to perceive the trajectory of a bullet shot from a gun, time must be slowed by a factor of 200,000. And the active components of electronic circuits take less than a nanosecond to switch on or off.

Molecules rotate in a matter of picoseconds. To image the vibrations of the chemical bonds between their atoms, one must increase the frame rate by another factor of 1,000 as each oscillation takes femtoseconds, in other words millionths of a billionth of a second (10^{-15} sec). To study the motions of the electrons in atoms or molecules, the shutter speed must be reduced to a few attoseconds. Attoworld's physicists routinely produce laser pulses in this range. And these ultrashort flashes make it possible to observe the dynamics of elementary particles in real time. For the first time, we have a direct window into a hitherto mysterious realm.



The goal is to understand and control the motion of electrons upon interaction with the electric light field of ultrashort pulses. The initial events take place as the electrons are being accelerated by the changing electric field of the incoming pulse on an attosecond time scale. This can then induce chemical follow-up reactions that may be relevant even for understanding biochemical processes in our body. The fast-changing electric field of ultrashort pulses can also be used to accelerate charged particles like electrons or protons. On this basis, light-driven particle accelerators are being developed which fit into standard-sized laboratories. Such accelerators are seen as a real alternative to the kilometer-sized and multi-billion-dollar costing facilities conventionally used for particle acceleration. The ability to control the behavior of electrons using light would also revolutionize the field of electronics. Ultraviolet light fields oscillate at rates around 1 PHz (10^{15} Hz). If electron flows could be switched at such a rate, opto-electronics would beat plain electronics by a factor of about 1,000.

The Center for Advanced Laser Applications (CALA) in Garching is now exploring the potential of light-accelerated electrons and protons for detecting and treating different types of cancer.

Laser science in healthcare and in medicine

Dr. Veit Ziegelmaier

When Theodore H. Maiman unveiled the first functional laser on July 7, 1960, in New York, it was hard to conceive of the range of applications that such high-intensity beams of tightly collimated light would one day find. But within a year, it made its debut as a surgical instrument in ophthalmology when it was tested on a patient in the US. And laser physicists have yet to run out of ideas and applications.

Medicine profits enormously from lasers. Lasers are used as high-precision scalpels. Unlike finely machined knives, the beam does not come into direct contact with the underlying tissue, making the device inherently sterile. A laser beam heats the skin and cuts it open or, if necessary, destroys it with pinpoint accuracy. There is no bleeding during treatment because the heat of the radiation immediately cauterizes the veins.

This also makes lasers suitable for stopping bleeding in a targeted way, even in places that are difficult to access. The laser beam reaches the affected area directly via endoscopes or light conductors and can thus stop internal bleeding in the stomach, for example. Laser light can also reach blood vessels directly via optical fibers and clear veins of unwanted deposits and blockages (arteriosclerosis) that could potentially lead to a heart attack.

Likewise, lasers can be used to remove superficial skin deformities such as port-wine stains or to erase tattoos by evaporating the pigments in the skin. They are also playing an increasing role in the diagnosis and treatment of tumors. Smaller tumors in the body that can be reached directly, such as skin cancer, can now be removed using lasers by killing the cancerous cells. Laser-guided computed tomography helps to obtain cross-sectional images of the body. The method is less dangerous and less costly than X-ray and nuclear spin procedures, which, however, still have a higher resolution than imaging methods.

One day, the use of ultrashort laser pulses could help to detect diseases at an early stage and thus provide more efficient treatments.

Pulsed lasers in the infrared radiation range have found an application in dentistry. Since the laser radiation heats up embedded water and as there is more water in a decayed tooth than in a healthy tooth, a tooth damaged in this way can be treated effectively without destroying the neighboring healthy teeth. In addition, it is possible to precisely scan the shape of a tooth using laser radiation. With the help of the data, a dental prosthesis can be produced optimally and precisely.

In ophthalmology, laser surgery is used to treat myopia, short-sightedness, and long-sightedness. Both conditions can be corrected by altering the curvature of the cornea, either by laser ablation of corneal substance (myopia) or by raising the level of the cornea (long-sightedness). As the uppermost corneal layer is not damaged, scar formation is avoided.

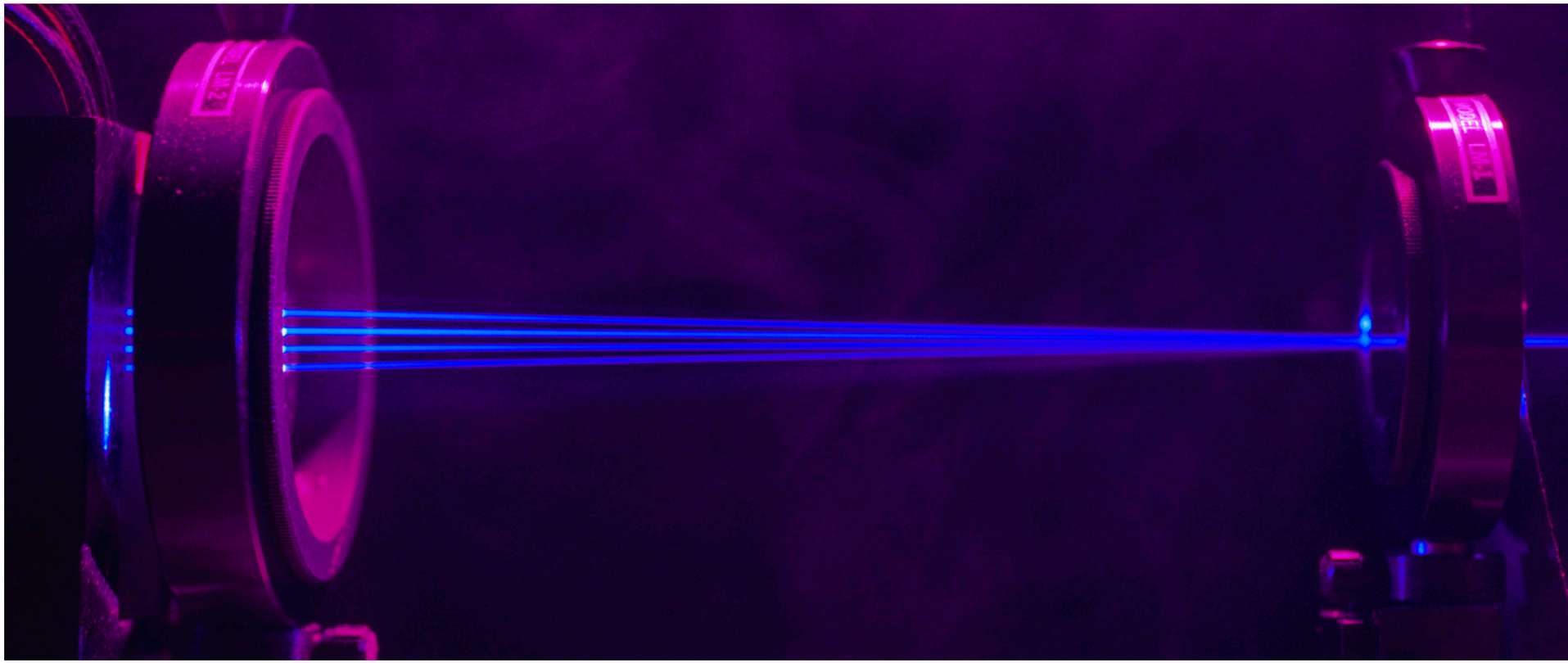
As we can see, lasers are already an indispensable part of medical technology. But the possibilities for the future are far from being exhausted. Just imagine: you make an appointment for a screening test. Your doctor takes a blood sample. The technician places the sample in a scanner and flips a switch. The sample is exposed to laser light. And fifteen minutes later you receive the result.

This scenario is still an aspiration. But it can become a reality. An interdisciplinary team of physicists, bioscientists, mathematicians and physicians based at the Center for Molecular Fingerprinting, in close cooperation with the Ludwig-Maximilians Universität München and the Max Planck Institute of Quantum Optics are now hard atwork on this project. Their goal is to develop a comprehensive blood test based on state-of-the-art laser technology. By using ultrashort pulses of laser light to analyze the molecules present in blood plasma samples, it should be possible to detect diseases such as cancer at an early stage.

To grow, divide and communicate with their neighbors, cells must synthesize, modify and break down molecules. These processes give rise to a complex array of organic compounds known as metabolites. Cancerous cells produce subtle differences in their metabolic products compared to those made by normal cells. Also, the body reacts differently to tumor cells, even if still small. Finally, all these substances end up in the bloodstream. But those that are specific for diseased states are few and difficult to detect. Laser technology could change that.

CMF researchers are working on identifying the characteristic molecular fingerprint of the non-cellular part of human blood as a proxy for defining healthy and diseased states. Depending on its structure, every molecule absorbs specific wavelengths of light. Laser pulses can therefore detect and measure the concentrations of the molecular mixtures present in the sample.

What sets this approach apart as being innovative? The researchers employ broadband infrared laser light, which enables them to quantify a wide range of biomolecules (carbohydrates, proteins and lipids) resulting in a comprehensive “global molecular fingerprint”. This new laser technology is more sensitive than current methods, with a finer nose for disorders.



Laser physics for health

Dr. Alex Weigel in conversation with Dr. Veit Ziegelmaier

“We are looking forward to our vision of a new type of healthcare” said Dr. Alexander Weigel, who leads the Laser Science Division at the Center for Molecular Fingerprinting. In his interview, Dr. Weigel gives an overview of the research activities and how laser research can be combined with engineering approaches to develop new technology for health monitoring using a novel technique called infrared molecular fingerprinting.

What exactly is your research about?

In the CMF's Laser Science team, we combine laser research with engineering approaches to develop new instruments for health monitoring using a novel technique called infrared molecular fingerprinting. The concept behind it is that we use an ultrashort infrared pulse to excite the molecules in a sample and then record the molecular response with a very sensitive detection technique. We develop our technology towards ever higher sensitivity and reproducibility. Together with Dr. Žigman's team, we are using our instruments in large-scale human blood-sample studies to identify infrared fingerprint signatures that indicate the possible development of diseases.

One day we aim to have an *in vitro* diagnostic test that can, based on a single drop of blood, be used to probe human health.

Could you explain in more detail what happens during a measurement?

It is similar to hitting a gong with a mallet and then listening to the decaying vibrations of the gong. In our case, we use our ultrashort infrared pulses as the mallet, and strike not a gong but molecules in a sample. We then "listen" to the electric field emitted by the vibrating molecules in response to the excitation. While the acoustic vibrations from a gong are easy to record, using a microphone for example, the electric field oscillations of light are much faster so that they cannot be directly measured with any available electronics. In order to measure them, we need to be faster – and the fastest that is available in nature is light itself. In reality, we use other ultrashort pulses that last only a few femtoseconds to scan the infrared response that is emitted. Just to give you a scale: a femtosecond is only a millionth of a billionth of a second. Returning to the picture of a gong being hit: You can imagine that the vibrating response depends on the gong's size, shape, and material. In the same way, the response of a molecule after being excited by an ultrashort infrared pulse depends largely on the size, shape, and the chemical bonds of the molecule. Therefore, we can use the infrared response to distinguish between different molecules. For health monitoring, we are interested in analyzing samples of human blood.

Why is blood such an interesting liquid?

Blood is a complex liquid and contains a huge variety of different molecules. Biologically relevant ones include proteins, carbohydrates, lipids and nucleic acids. With our technique, we measure the complex vibrational response signal of the blood sample, which we call its molecular fingerprint. It contains response contributions of all the molecules in the sample. While we may not be able to distinguish each individual molecule, the fingerprint signal still represents the complex composition of the blood sample and is very sensitive to changes in molecular composition.

What is the vision behind it?

Our vision is to identify characteristic signatures in the fingerprints of blood samples that are specific to human health. Human blood is the fluid that interconnects all parts in the human body. Therefore, when a disease evolves, we expect small changes to appear in the molecular composition of the blood sample. The idea is that if we are sensitive enough to detect these changes in the molecular fingerprint signal, we can possibly diagnose any medically relevant deviations from a healthy state using a simple blood-based measurement. CMF has already started a large scale clinical study, called Health for Hungary - Hungary for Health (H4H) involving thousands of individuals whose blood samples will be investigated over a number of years using the technique we are developing.

What are you working on in particular?

The task of my team is to provide the technology and develop the instruments to make this vision a reality. First of all, we need powerful and extremely short infrared lasers. One day we aim to have an *in vitro* diagnostic test that can, based on a single drop of blood, be used to probe human health. First of all, we need powerful and extremely short infrared lasers. Part of my team is constantly pushing the limits of what is possible with today's laser technology and is developing new powerful femtosecond infrared sources. On the other side, we are also working on the research and development of new detection solutions, not only on the optical and electronics side but also on the data processing side, to reach ever-higher levels of detection sensitivity. Our engineers are combining our research results into ultra-stable and user-friendly instruments to measure and compare human-blood samples from large-scale campaigns over many years.

If we can do this, it will be a game changer in the field of in vitro diagnostics: Being able to do complex health monitoring based on a drop of blood.

What are your next research goals?

On the one hand, we are constantly striving to improve our laser technology, which is the foundation for blood-based assay development. Our next-generation instrument is based on Cr:ZnS laser technology. This new type of ultrafast laser is particularly suitable for generating infrared pulses. The exact shape of the infrared pulses that excite the sample also influences the molecular response signal we measure. It is incredibly important to create reproducible infrared pulses from the laser. So we actively stabilize our laser systems to maintain the full waveform of the pulses exactly over and over again. Naturally, our detection needs to keep pace with laser development: we need to find ways to take measurements with the same spectral bandwidth that our laser pulses provide. Not only that: we want to record in a single measurement the strong infrared excitation pulse simultaneously with ever weaker response signals. The property of the largest measurable signal versus the smallest detectable signal is called dynamic range. Achieving an extremely high dynamic range requires both research and development on the optical side of detection, and the development of new electronic detector solutions on the other side. The ultra-low-noise performance of our lasers is another crucial parameter to achieve the highest sensitivity in our measurements. Any noise in our laser output directly translates into measurement noise, which can hinder us from seeing small changes in the molecular fingerprint response. In our latest-generation lasers, we use pumping with telecommunication-grade semiconductor laser diodes as one of the key technologies to achieve extremely low fluctuations of the laser. In our instruments, we are synchronizing two lasers to perform extremely fast measurements. These are so fast that during a single scan many of the residual noise contributions are frozen in time. We are looking forward to applying the new instruments for molecular fingerprinting measurements on the new samples from the H4H Program. The goal is to identify, together with data analysts, specific molecular fingerprint signatures that report on the development of human health and the possible development of diseases like cancer.

When do you think your method will one day find its way into medical practice?

Last year, as part of the Lasers4Life clinical study at LMU, we carried out a measurement campaign in which we investigated blood plasma from more than 5,000 human individuals, and next we will perform systematic measurements on the samples of the H4H Program. Still, it is a long way from taking first measurements to having instruments used for regular medical diagnosis. We are talking here about years of hard work ahead of us, plus clinical trials, etc. On the measurement side, we first need to evaluate our technique on many blood samples and learn how to take into account, for example, the diversity of the human population and individual lifestyles and their influence on our measurements. CMF is currently establishing new facilities in Szeged to store and measure a large number of human blood samples under controlled conditions. There is a lot of research ahead of us in the upcoming years before we can establish our technique as a regular diagnosis tool, and my team and I and all our collaborators are working together towards making this vision reality.

Infrared molecular fingerprinting

Changes in the body's health state cause characteristic changes in the molecular composition of human blood. When a drop of blood plasma is exposed to ultrashort laser light, the light waves emitted by the excited molecules of the sample are directly detected with femtosecond-attosecond laser techniques. Electric field molecular fingerprinting (EMF) has the potential to capture even minor differences in the molecular composition of human blood, therefore indicating changes in the individual's health state. CMF is driven by this vision to define the frontiers of probing human health based on the technological, ultrashort pulsed laser development of molecular fingerprinting for the next generation of molecular diagnostics.

Infrared molecular fingerprinting is an *in vitro* diagnostic method based on pulsed laser technology. A drop of blood plasma is exposed to ultrashort laser light, and the light waves emitted by the excited molecules of the sample are directly detected with femtosecond-attosecond laser techniques.

The molecular composition of systemic biofluids, such as our blood, can be used as an indicator of human physiological states, which is useful for disease detection. The ability to observe minuscule changes in the concentration of molecular signatures of liquid biopsies (blood plasma and serum) makes it possible to advance medical diagnostics.

Molecules of non-cellular parts of blood samples are illuminated by ultrashort laser pulses. The abruptly excited molecules emit light at characteristic frequencies in the wake of the excitation. The electric field of this molecular signal is detected with a femtosecond-attosecond resolution using latest ultrafast laser technology. The new approach – field-resolved laser molecular fingerprinting – is well suited for detecting changes in the concentration of a wide variety of molecular types (proteins, lipids, carbohydrates and amino acids) within a single measurement. Infrared molecular fingerprints are promising for improving the specificity and sensitivity of disease detection, possibly in combination with other biomarking approaches.

Zooming into the molecular zoo of our blood

Dr. Mihaela Žigman

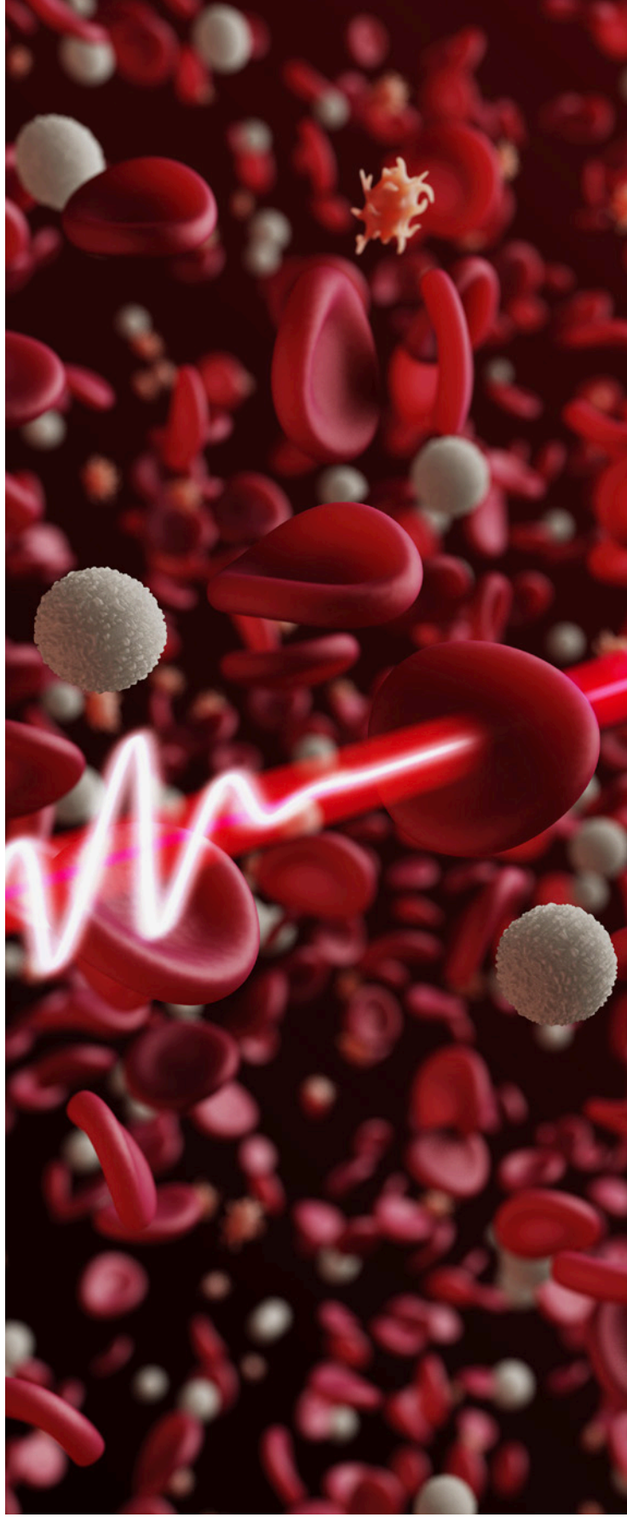
The pursuit of understanding the human body's intricate biochemistry over the centuries has contributed to our current level of wellbeing as a species. One crucial aspect of this understanding lies in comprehending the molecules present in human blood, which has shaped modern medicine. In the late 1800s, scientists made pivotal discoveries about blood components such as glucose, urea, and proteins, laying the groundwork for clinical chemistry testing and its potential application in disease diagnosis. Today, the quest is for advanced molecular methods that will enable early disease detection and personalized medical treatments tailored to the characteristics of individuals.

At CMF, we are focused on developing new ways of blood-based testing, based on our advancements in ultrashort laser technologies at our labs. Blood testing is a fundamental minimally invasive diagnostic approach in healthcare. And our goal is to make it even more effective. We measure the non-cellular part of human blood, as taken by your doctor, anytime you go for a check-up, with infrared spectroscopy. We analyze the specific vibrations of molecules present in blood plasma by shining laser light pulses on them. The concept is conceptually comparable to tapping a tuning fork. After we bang on it, we keep quiet and listen to the signals from the bang using the very sensitive technology we are developing. It is as if we were listening to and recording soundwaves emerging from the molecules that we shined the light on. To be in a position to "hear" and tell apart the tiniest differences (if the recorded sequences differ from the expected "sounds") we use state-of-the-art computational tools. We build algorithms to distinguish the person's healthy and unhealthy states using just a small blood sample.

We aim to gain a deeper understanding of the molecular mixtures in our blood to catch disease early and enhance patient outcomes and overall quality of life.

Each person is unique and undergoes different changes over time, particularly as we age. This variability poses a significant challenge in medicine, as it is difficult to identify the precise moment when a person starts developing a treatable health disorder. To address this challenge, we have designed the H4H Program. Within this framework, we follow individuals over time and ask them to donate a bit of their blood and medical information, all anonymously. We aim to gain a deeper understanding of the molecular mixtures in our blood to catch disease early and enhance patient outcomes and overall quality of life. Developing innovative procedures that do not yet exist and establishing them according to all relevant scientific standards is a long and challenging road. We are actively working towards this goal. With over 1,000 volunteers having returned to our centers across Hungary on more than four occasions, we are on a positive path. And we want to go even further. We wish to learn more about how our blood changes over months and years to be able to better tell whether there is a likelihood that ill health is on the horizon. And the more people that join in the better, and the sooner we can do this. We are already seeing our research having an effect, with several participants becoming aware of previously undetected conditions through conventional blood tests, leading to timely treatment for issues like abnormal cholesterol levels and blood pressure. These seemingly minor concerns could have escalated into more severe health problems if left unchecked. Are we close to "hearing" and distinguishing medically relevant molecular compositions of the molecular zoo dissolved in our blood? We are currently setting the foundations. Anyone involved is aware that it is not a single research group endeavor. There is a whole research program umbrella around it, and professionals from diverse research fields are working within it. But even this is not enough: we also need the contribution from the individuals that agreed to join in and help this research program to thrive. While it will take several more years to develop and make an infrared laser light medical blood test available, we have already made a substantial contribution to the health of the Hungarian population. As our laser technologies continue to advance and the H4H Program gets more participants, the blood testing methods developed at CMF will gain ever greater value, unlocking new possibilities for medical diagnostics. The knowledge and experience we are gaining will empower individuals to take control of their health, ultimately contributing to an improved quality of life and well-being. In our work, we are finding new ways to disentangle patterns of molecules in blood that have not been captured before.

Decoding the patterns observed in the behavior of molecules in living systems has been a remarkable scientific journey. This knowledge has transformed various scientific disciplines, from chemistry and physics to biology and medicine, opening up new possibilities for technological advancements and improving our quality of life. Infrared spectroscopic fingerprinting comes in as a new player, as it has not been applied for monitoring human health. With our valiant zeal, we are not only introducing this novel approach but are also likely to be turning a new page here.



The H4H Program: Health for Hungary – Hungary for Health Clinical Study

Dr. Zoltán Kovács,
Dr. Domokos Gerő

The H4H (Health for Hungary - Hungary for Health) Program is part of a novel research program in which we are collecting blood samples from an estimated fifteen thousand healthy volunteers over 10 years. Our goal is to develop and verify a new way of measuring human health – an approach called infrared molecular fingerprinting. We aim to establish a way of analyzing human blood with newly developed laser-based technology, characterize and monitor health, and detect possible early signs of developing diseases. Individuals will be followed over time to be able to measure and monitor the onset of possible diseases in a sensitive way, with personalized differences being measured. Once our approach has been verified, the H4H Program will be able to contribute to creating a reliable health monitoring system that will improve quality of life for all of us.

How does the clinical research work?

Thanks to the advancement of medicine and technology, many diseases can now be diagnosed and treated very effectively, thus lengthening the time people spend in good health. However, there is a well-defined need for continuous development of even more specific and sensitive diagnostic methods and therapies; and as in sports, there is always room for improvement, especially against the backdrop of an ageing population.

In the history of humankind, there are several written sources of controlled clinical-like trials, the first of them being probably the experiment of the ruler of Babylon, King Nebuchadnezzar, somewhere around 600 BC. This is described in the Book of Daniel in The Bible[1], when due to opposition of some Israeli captives to eat non-kosher food a 10-day trial was carried out which proved that a diet of legumes and water was just as good for you as eating meat and wine (which was at the time considered the best way to stay in good physical condition). As a result, the captives could stick to their meal preferences.

A long road from this story leads to today's GCP (Good Clinical Practice) era, in which clinical trials are being performed based on standardized rules and ethical considerations. All medications and medical devices must undergo a series of preclinical studies before being tested on humans. They can then be tested on humans in clinical trials. Only once the various phases of clinical research have been successfully completed can any medication or device be marketed, and then only for the purpose it has been proven to be useful for.

There are several key participants in each clinical trial, including the Sponsor of the trial, who has an idea they want to realize; the regulatory authorities and ethics committees, who safeguard the rights of the subjects and future patients; the investigators together with their designated investigational team of doctors, nurses, and coordinators, who perform the trials at their investigational centers with outreach to appropriate subjects and necessary logistics; and most importantly the volunteers, who are either healthy or diseased subjects who offer their time and bodies to undergo the procedures described in the trial protocol. A team of clinical research professionals and their operative support helps to get everything done according to legislation, rules, and protocol. They also make sure all procedures and documents are verified so the data that has been collected can lead to a scientific conclusion.

It is crucial that a successful clinical trial is not aimed at leading to a definitively positive result for Sponsor. Rather, it needs to prove a viable idea that it can be used clinically, but no less importantly, if the idea is not viable, it must be allowed to lead to a negative conclusion, or suggest necessary changes.

---

[1] *Collier R. Legumes, lemons and streptomycin: A short history of the clinical trial. CMAJ. 2009;180:23–24*

What is the rationale of the H4H Program?

In 2021, CMF started its first clinical research study in Hungary, called the Health for Hungary – Hungary for Health Clinical Study. It is a longitudinal clinical study specifically designed to promote the applications of molecular fingerprinting into clinical translation, particularly for population health monitoring.

In our society, a large portion of deaths are attributable to preventable diseases, and these diseases also cause the people affected by them to become limited physically and even disabled. Early diagnosis and proper treatment of conditions that are precursors to serious illness are believed to serve as major contributors to prevention strategies and help reduce the burden of morbidity. However, most available medical tests are not ideally suited for health monitoring: they were developed to detect symptomatic diseases and have been secondarily adapted for applications in health monitoring: to follow up healthy individuals with the aim of recognizing the early signs of abnormalities in pre-symptomatic subjects.

Electric field-resolved molecular fingerprinting (EMF) represents a new technological platform that allows the detection of molecules in minute concentrations that was not possible in the past. The detection sensitivity of the methodology, in combination with the speed at which the measurement device collects data, opens up new horizons in the field of medical diagnostics and *in vitro* blood tests, and may serve as a valuable tool for health monitoring in the future.

The H4H Program is a prerequisite for streamlining the move of this new technology into clinical use. Since EMF and the way we analyze the data we collect presents a new approach, a normal range of measured infrared fingerprints and their possible deviations over time needs to be established before it can be used for detecting diseases in practice. Once a reference range has been established, deviation from the normal healthy state may serve as an indicator of health abnormalities and a sign of early-stage disease, even in the absence of clinical symptoms.

What are the goals of the H4H Program?

In order to develop EMF as a novel *in vitro* diagnostic tool, from concept to testing its clinical utility, in the H4H Program we are pursuing the following goals:

1) Establish the healthy baseline of blood-based infrared molecular fingerprints.

2) Establish clinical decision limits of EMF that are diagnostic for early-stage development of select non-communicable diseases (NCDs), including lung cancer, cardiovascular disease (CVD) and diabetes mellitus.

Ad. 1. A reference interval is usually defined as the 95% confidence interval of measurement values for a healthy (reference) population, which is based on the between-subject variability of laboratory test results. As blood plasma constituents are known to vary with age, gender and other parameters such as smoking and obesity (assessed as body mass index, BMI), we foresee specific reference ranges for each of these healthy adult sub-populations (e.g., age-group specific reference intervals). Alternatively, temporal trajectories of EMF will be determined as normal patterns that develop with age and that are typical of the predefined subpopulations. These allow the prediction of expected values for a given age.

We also presume that a person-specific normal range of EMF is of considerable practical use, reflecting very little variation in individuals over time. The low within-subject variability may allow for an even narrower, individualized normal range that may prove more appropriate for detecting early-stage health deviations.

Establishing a reference range of infrared molecular fingerprints and its deviations, based on a healthy population, may initially seem of limited clinical use, though, as it is very much specific to health, and this information may not be directly transposed to the likelihood of disease: test results outside the reference range may not indicate a disease. Yet, it will have significant impact for the second milestone.

Ad. 2. The ultimate goal of novel diagnostics is to aid clinical decision making. Therefore, it is important to establish clinical decision limits: diagnostic results

BASIC STUDY DATA

H4H TOGETHER FOR HEALTH

STUDY DESIGN
› SINGLE COUNTRY
› MULTI-CENTRIC
› PROSPECTIVE TRIAL

TOTAL LENGTH OF THE STUDY
10 YEARS

TWO STUDY ARMS

1. Low-and moderate risk study arm
2. High-risk study arm

NUMBER OF INVESTIGATION CENTERS **20+**

TOTAL NUMBER OF PARTICIPANTS **15,000**

ENROLLMENT START July 27, 2021 › › › **PLANNED COMPLETION DATE** December 31, 2030

MEASUREMENTS AND MILESTONES (DECEMBER 2023)

NUMBER OF PARTICIPANTS ENROLLED	NUMBER OF VISITS AND BLOOD TESTS COMPLETED	NUMBER OF BLOOD SAMPLE MEASUREMENTS
10,522	**40,519**	**12,860**

HEALTH ASSESSMENT METHODOLOGY

DETAILED HEALTH QUESTIONNAIRE	PAST MEDICAL HISTORY AND MEDICATIONS
CLINICAL LABORATORY TESTS	ROUTINE BLOOD TEST PANEL
MEDICAL TESTS IN HIGH-RISK ARM	BODY COMPOSITION, ELECTROCARDIOGRAM (ECG)
	LOW-DOSE CHEST CT, CORONARY ARTERY CALCIUM SCORE

criteria that can be used in clinical medicine to make decisions, i.e., proceed with diagnostic tests (e.g., imaging methods), decide on an intervention, or choose a treatment modality. This aim will be best fulfilled if the criteria are supported by evidence of disease or are based on clinical outcomes. If a symptomatic clinical disease is present, selecting individuals with a well-defined condition, specific outcomes may be chosen for setting up clinical decision limits (e.g., diabetic individuals are followed up for the development of complications). However, the association between specific medical test results and the presence of disease is more difficult if symptomatic disease develops years later in seemingly healthy people. For these conditions (i.e., lung cancer and CVD) follow-up is needed until the diagnosis is given.

In the framework of the H4H Program, to verify the clinical validity of EMF results, individuals are followed for up to 10 years. Blood plasma collection and analysis are continued throughout the observation period and the development of clinical disease is noted to precisely record the time of diagnosis to help validate EMF data.

How is the H4H Program organized to pursue these goals?

In the framework of the H4H Program, we collect blood samples from 15,000 healthy volunteers over a 10-year period. To achieve these objectives, the H4H Program is running at more than twenty investigational centers under the guidance of CMF, with the involvement of 5-15 team members at each of the centers. A lot of healthy participants are required to set up age-group and gender-specific reference ranges, and a long follow-up period is necessary to observe the development of diseases, due to the relatively low incidence of events.

Participants are enrolled into two arms of the study: (1) a low- and moderate-risk cohort and (2) a high-risk cohort of subjects, corresponding to the risk of cardiovascular disease development. In the low- and moderate-risk cohort follow-up is foreseen for ten years, while the sample collection period is expected to run for five years in high-risk individuals. The two cohorts are expected to provide complimentary information: the low- and moderate-risk cohort will be superior in establishing the population reference intervals, and the high-risk cohort will provide the bulk of disease-specific signals. Personalized reference ranges will be determined in all participants, and disease development is expected in both cohorts. We expect that approximately 400 fatal or non-fatal CVD events may occur in the low- and moderate-risk cohort over ten years, and around twenty five cases of lung cancer will be detected. The high-risk cohort participants are older subjects and thus the risk of lung cancer is considerably higher. Therefore, hundred newly diagnosed lung cancer cases are foreseen on this arm of the study within five years. Also, the relative frequency of CVD will be

higher on this study arm but this is counterbalanced by the shorter observation period. Type 2 diabetes mellitus prevalence is around 10% in subjects enrolled in the study thus far, allowing us to study the actual development of diabetic complications with sufficient prevalence for subsequent statistical analysis over the given observation period.

All subjects are followed up with between one and three visits a year, during which health data is collected, clinical laboratory blood tests are performed, and plasma samples are collected for EMF measurements. Samples are transferred to storage temperature (-80 °C) within three hours of blood draw and are transferred to CMF's designated biorepository for long-term cryopreservation. High-risk individuals will undergo additional medical imaging tests and medical screens to check for cardiovascular disease and lung cancer to verify the health status of subjects.

What makes the H4H Program unique both in Hungary and internationally?

The H4H Program is a cutting-edge research study with a longitudinal prospective design that is not comparable to any prior clinical trials conducted in Hungary before. As part of this daring project, we have built up a trusted network consortium of clinical centers and involved more than 10,000 subjects in the study in less than a year and a half, with the prospect of further 5,000 participants to be recruited in the upcoming year. Regarding the scale of the clinical study, this number is near equal to the total number of subjects being enrolled in all therapeutic interventional clinical trials ongoing across Hungary; and it is highly representative of the population, also functioning as an observational study: corresponding to 0.2-0.3% of all middle-aged and elderly people in Hungary. Since all health data is carefully checked by clinical monitors, and any incongruences are corrected according to the original clinical records, the data quality of H4H Program is superior to typical observational studies, showing a close resemblance to interventional drug trials in this respect.

Moreover, the H4H Program is an exceptional endeavor due to several factors all over the world, and not only in Hungary. The study aims to facilitate the clinical translation of a novel health monitoring medical test that is not based on disease-specific markers but on the entirety of molecules dissolved in human blood plasma, as a biomarker-agnostic, holistic approach. EMF allows the detection and analysis of unknown constituents of the plasma, unlike other methods that only focus on already known proteins, nucleic acids or metabolites. This means it may be able to identify novel tumor-associated fusion-proteins that could change the future of diagnostic medicine.

While the approach detailed above is expected to provide positive results, the study incorporates two inherent fallback strategies, which sets it apart from similar studies: (1) the possibility for retrospective (re-) analysis of samples and (2) an option for expanding the study with additional high-content diagnostics. As opposed to the prospective design, the former approach lets us re-evaluate the disease-specific spectral patterns and makes it possible to identify additional spectral markers after the respective conditions have been diagnosed. The second option is a feasible expansion of the study to optimize early disease detection: cryopreserved plasma may be re-analyzed in the future either by a more advanced molecular fingerprint analyzer or using any omic methods (e.g., mass spectrometry-based proteomics, glycomics or metabolomics) that supplement the acquired data to provide a more complete health assessment methodology.

The study aims to facilitate the clinical translation of a novel health monitoring medical test that is not based on disease-specific markers but on the entirety of molecules dissolved in human blood plasma, as a biomarker-agnostic, holistic approach.

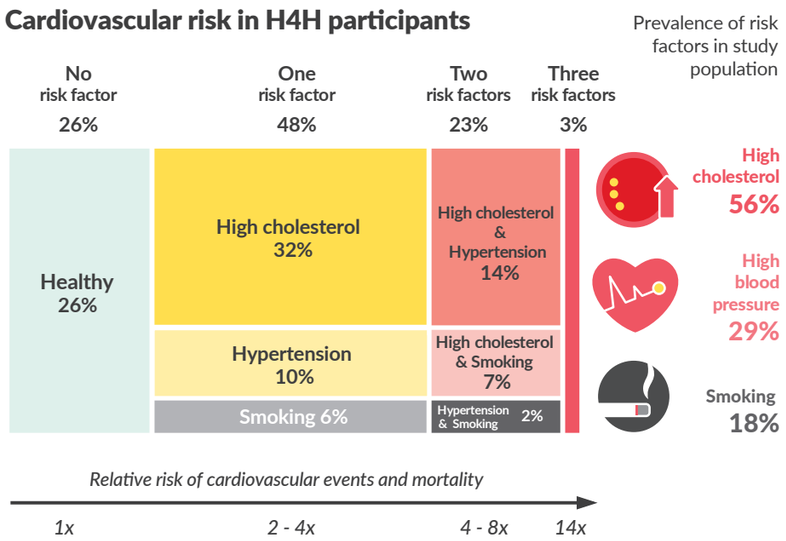
Overall, the H4H Program has already delivered some very positive effects for the participants, but there is still a long way to go to achieve the final goals of the program. Given the recurrent routine laboratory testing and regular check-ups by medical professionals, many subjects have been made aware of several already underlying conditions they did not know about. Also, the risk of disease was uncovered in some subjects in cases in which a condition was not adequately suppressed by the applied therapy (e.g., high cholesterol level despite receiving lipid-lowering treatment). As a result, treatment initiation or adjustments were recommended for some of the enrolled subjects that will potentially lead to longer life expectancy and a better life quality.

Preventive medicine and personalized healthcare

Dr. Domokos Gerő

CMF believes that providing personalized healthcare tailored to each person's unique biology, and identifying diseases before the symptoms arise is the future. Our research could make the current healthcare system more effective and improve the quality of preventive medicine, with the possibility of safeguarding our health in the future.

Cardiovascular risk factors in participants of the H4H Program. The prevalence of the three main risk factors (hypertension, hypercholesterolemia and smoking) is shown. Box sizes are proportional to the relative frequency of conditions and their co-existence in the study population, and compared to the number of healthy individuals (subjects with no risk factor). The associated risk of cardiovascular events and death rates are shown on the bottom.



What is preventive medicine?

Medicine was developed to provide help with illnesses and to treat health problems that compromise a person's well-being. As it was soon realized, especially with epidemics, preventing the spread of disease was easier than treating a fully developed condition. Thus, preventive medicine was conceptualized to forego disease occurrence and halt the progression of clinical conditions by applying measures in advance.

Controlling communicable diseases has been especially effective, and during the 20th century, these efforts radically decreased the number of premature deaths. Over recent decades, however, significant changes have taken place, as from the mid-1990s, the burden of non-communicable diseases (NCDs) has vastly surpassed the impact of infections. NCDs, which comprise two-thirds of the global disease burden, are responsible for 70% of mortalities worldwide, while in Hungary, 94% of all deaths are caused by NCDs.

Four main NCDs account for the vast majority of premature deaths: cardiovascular disease, cancer, diabetes, and chronic respiratory disease. So these four are the major targets of preventive care. Prevention strategies aim to stop disease development at an incipient stage, as early as the risk becomes evident. So risk factors (conditions predisposing to disease) are identified before they cause any symptoms. Clinical screening for risk factors occurs in healthy individuals in the context of primary care. Among the NCDs, cardiovascular disease (CVD) is the most prevalent, and it is widely believed to be largely preventable: 80% of CVD deaths might be avoided by early detection and proper management of risk factors. The interventions include a combination of lifestyle changes (dietary changes, promotion of physical activity and weight loss) and pharmacological therapy. The application of appropriate treatment, which helps prevent the development of clinical disease, is known as primary prevention, as it focuses on preventing diseases from occurring.

Primary prevention is less effective in Hungary than in Western European countries, and this issue is the main cause of the shorter healthy lifespan in the country.

What is personalized medicine?

While a disease is believed to indicate a homogenous entity caused by a unique underlying process, it may cause a wide range of symptoms that also vary in severity and pattern in each individual. Similarly, disease development is not a unified process: in a group of people who apparently share the same characteristics of predisposition, some present earlier with clinical disease than others. Furthermore, disease progression also shows a high degree of variability in people. In this case, as it seems obvious, the one size fits all approach will not suffice, as it cannot provide equal benefit to all people. Instead, a strategy is needed to reduce the risk of disease with the use of differentiated care in individuals. This is achieved by tailor-made prevention steps and individualized therapy, using stratification to set up subgroups of people with similar risk levels and by applying a treatment plan based on the risk category. An even more sophisticated approach is foreseen in the future, with further individualization, providing personal care based on a more detailed risk assessment. These strategies, collectively known as personalized medicine are necessary to extend healthy life further and achieve better health outcomes in society.

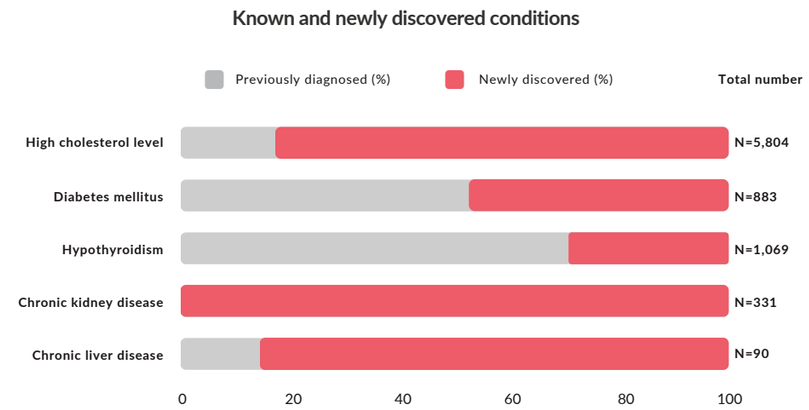
What does the H4H Program do to promote prevention in healthcare?

The health data that is collected using questionnaires and routine blood tests as part of the H4H Program is necessary to confirm the absence of disease, to perform risk stratification, and to allow the monitoring of risk factors. The consistency of the data that is collected is also monitored. With regards to some parameters, it allows us to identify conditions that were overlooked, and either lack guideline-directed medical therapy or were missed completely in the clinical routine. When such conditions are noted, individuals are referred to their primary care physician to initiate proper management of the abnormality. It is essential to take notice of these conditions, as they rarely cause complaints, but they silently promote diseases.

The most common conditions that were underrecognized included high cholesterol levels, diabetes, and thyroid disease. Unfortunately, all of these increase the risk of cardiovascular diseases and shorten life expectancy if they are not managed properly. During the follow-up of subjects, in the context of the program, blood tests are repeated to check the effectiveness of the applied treatment. Undertreatment of risk factors is, unfortunately, a common practice around the world, and it has also frequently been seen in participants in the H4H Program. If a condition does not cause any complaints and the individual receives treatment, it may give a false impression that everything is fine. However, therapeutic goals must be achieved to reduce the risk. These goals are often expressed as laboratory values that are known to be associated with reduced risk levels. If the recommended target values are missed in participants, investigation center personnel will recommend seeking medical advice. We estimate that thanks to timely recognition of conditions and early treatment initiation, H4H Program participants may experience a longer period of healthy life.

How shall the H4H Program help with health surveillance in the future?

The long-term goal of the H4H Program is to develop a novel tool for monitoring the health of the entire population. In contrast to current methodologies designed to screen for a limited number of diseases, our approach is distinctly different: we aim to monitor the health of every individual by analyzing a broad range of blood constituents. To identify deviations from the norm, we compare measurement results to an individualized, person-specific normal range. Ideally, this strategy allows for the earlier recognition of abnormalities and provides a more comprehensive coverage of diseases. We anticipate that laser-based, population-wide screening will contribute to preventing many premature deaths and increasing life expectancy.

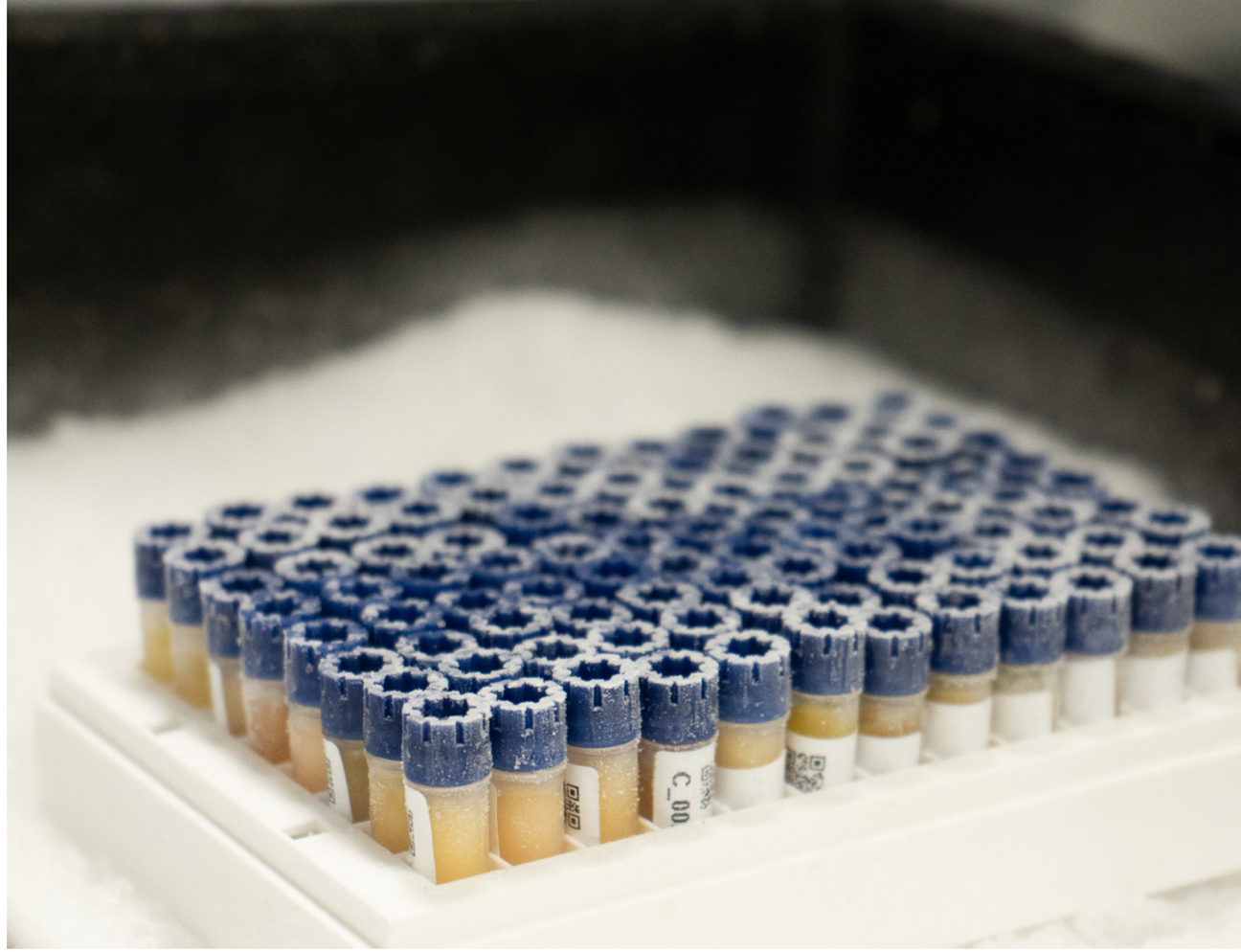


Newly discovered conditions in study participants. Common conditions and diseases are shown with the number of cases. Previously diagnosed conditions are shown on the left (grey) and the newly recognized ones on the right (red) based on their relative frequency. Many cases were missed earlier in the clinical routine but have been discovered with the help of the study.

The hidden challenges of freezing and storing biological samples

Dr. Frank Fleischmann, Dr. Diána Debreceni

During our multi-centric H4H Clinical Study, we aim to store hundreds of thousands of blood plasma samples collected at twenty different collaborating institutions (study sites). These samples are collected using pseudonyms and adhere to biobank standards. CMF obtained its biobank license in July 2022 when the National Public Health Center (NNK) granted the license for its interim biobank in Szeged, which has a storage capacity of 80,000 aliquots.



Why not all ice is the same, and what this has to do with CMF

A once juicy chicken breast has been hiding at the very back of the freezer. But after reappearing almost two years later it has changed. Ice crystals have formed inside the freezer bag, and the meat has become dull and dry in some places. This is freezer burn. And this has happened despite the fact that it has been at a constant -25 °C in the freezer.

If you think that all ice is alike you are mistaken! Frozen water, is a dynamic substance. New crystals are constantly forming, shifting, and getting bigger, and ice can even evaporate.

The same thing that makes food go bad in a household freezer is great cause for concern at biobanks. Their job is to store biological samples, mostly intended for medical research. They often have to store these samples for years in such a way that the quality never deteriorates. In the H4H Program, for example, blood is drawn from volunteers two or three times a year for five years, maybe even longer, and the plasma is frozen for later examination. So, the first blood sample is already five years old when the last one has just been collected. And the researchers want to compare these with each other without quality issues.

Unlike meat, blood plasma is a liquid that contains no cells or tissue. And yet the makeup of blood plasma also changes over time under the influence of the ice crystals. If we look at what plasma is actually made of, almost 90% of it is water. Of the remaining 10%, just over half comprises proteins. These are long-chained biomolecules that can consist of several hundred amino acids. In order to perform their function, these amino acid chains are folded or crumpled together based on certain rules. Proteins can be over 10,000 times heavier than a water molecule. And the small water molecules also get into the cavities of the protein structure. If several water molecules come together to form crystals during freezing, this can become a problem for large molecules like a protein, especially since ice takes up a larger volume than liquid water. It is impossible to avoid ice crystals forming during freezing and damaging even a very fresh blood sample. But fortunately, biobanks have an ally that prevents ice crystals from growing. And that is the cold. The lower the temperature, the longer it takes for the ice to recrystallize. That's why most bio-laboratories have ultra-low temperature freezers where samples are stored at -80 °C. But even in conditions that are considered cold even for the Arctic, ice can still change, albeit much more slowly than in a household freezer. However, to store plasma for years, maybe even decades, even -80 °C is still too high.

But fortunately, even the ice crystals eventually get too cold. At around -135 °C and below, the crystals do not change any further. According to current knowledge, plasma samples can be preserved for decades under these cryogenic conditions without any noticeable change in quality.

Creating an environment this cold is not easy. Although there are very sophisticated special freezers that work similarly to our household devices and can reach such low temperatures, they consume a lot of energy. Therefore, a very cold liquid, specifically liquefied nitrogen, is often used as a coolant. Nitrogen gas makes up 78% of our ambient air and it's boiling point is -196 °C, which means it has to be colder than this to remain in its liquid form. The samples do not float in this liquid but are actually in the cold gas phase above it. There, the temperature ranges from a still very frosty -185 °C directly above the liquid nitrogen to -160 °C in the upper level of the storage room.

In the gas above the liquid nitrogen, samples can be stored safely for decades. But that only makes sense if you want to use the samples at a later date, which is when the next problem arises. The samples are not all stored individually in the biobank but in large sample racks. If you want to retrieve just a single tube from storage, you always have to take out a whole rack with dozens or even hundreds of tubes. For those samples that are not retrieved, it means physical stress when they are suddenly taken out of their very cold environment and later brought back to it. It is crucial to make sure that the samples do not thaw. This can be done by storing them on dry ice at -78.5 °C, but this still means a temperature shock of up to 100 °C for the samples, which is like throwing an ice cube into boiling water. This leads to stress cracks in the sample, and if the procedure is repeated, the quality of the samples also suffers. So it is important to try to sort out the samples at as low a temperature as possible. It goes without saying that this cannot be done by hand, and it is a very technical challenge. Some companies have succeeded in developing systems in which samples are automatically taken from the liquid nitrogen storage and sorted out at -130 °C. This is a technical feat and probably the limit of what is currently possible. Although the samples still suffer a temperature shock of up to 50 °C, there is currently no better way to do it.

A small, semi-automatic version of this kind of system is already being used by our colleagues at the Ludwig-Maximilians-Universität München (LMU). The long term goal for CMF's biobank is to build up a fully automated storage system, possibly based on the above-mentioned technology, that can cope with the large sample numbers of the H4H Program. With such an outstanding biobank, CMF would be able to test each new generation of laser spectrometers on optimally stored samples.

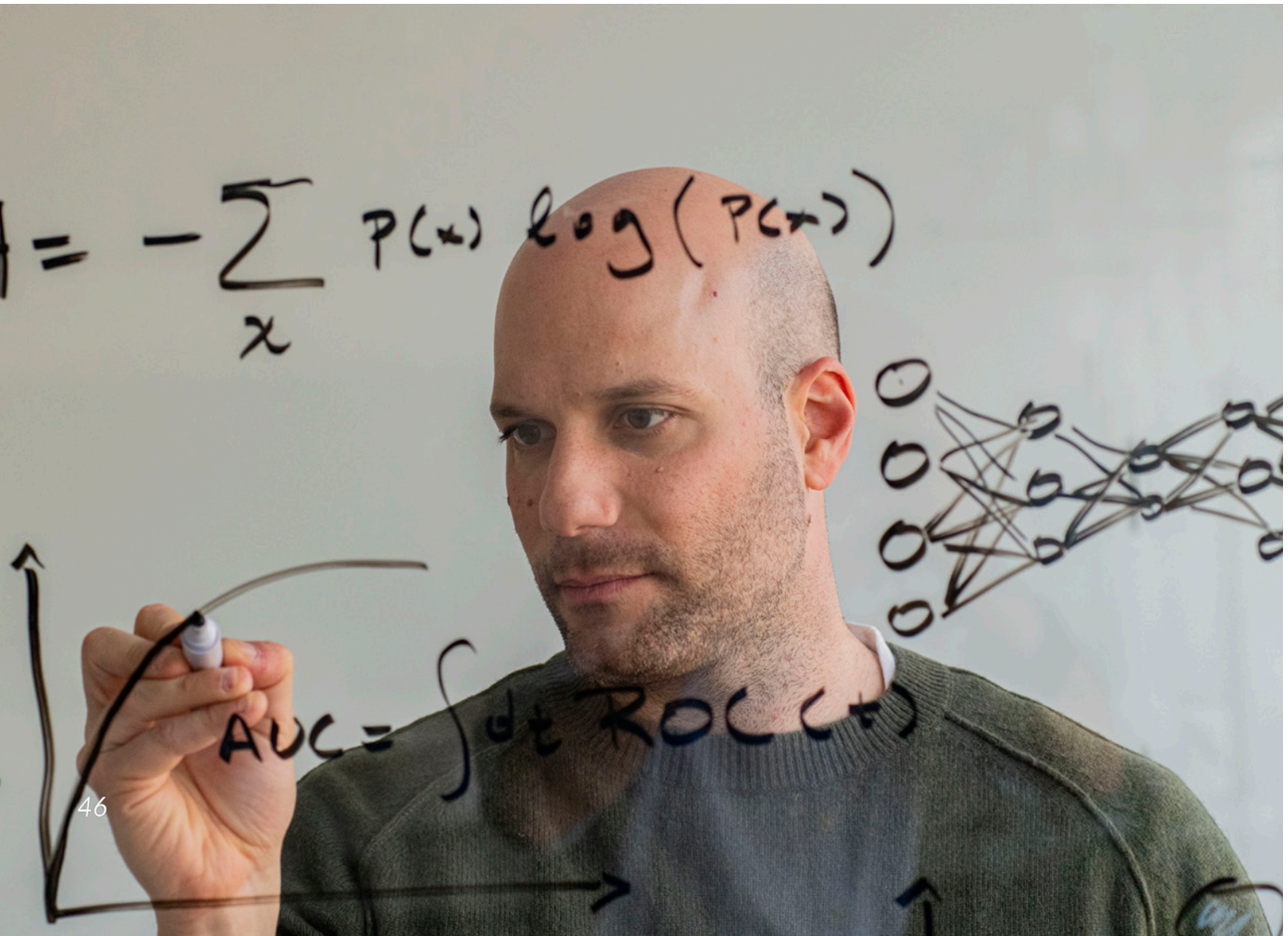
How can we distinguish so many samples?

The samples are labeled with a unique 1D alphanumeric identifier. Each aliquot of the same samples (10 per sample) is sorted into biobanking tubes individually marked with 2D-identification codes (inverse datamatrix codes on the bottom), which allow accurate biobank cataloging of the tubes. Clinical data associated with the samples is obtained from an electronic clinical database (electronic case report form (eCRF) platform) which is also assigned to sample IDs in a pseudonymized manner. For the traceability of the storage position and condition of the samples and aliquots, as well as the clinical and measurement data, a sophisticated but user-friendly modular Laboratory Information Management System (LIMS) is essential for the daily operation of our facilities.

The role of Data Science

Dr. Kosmas Kepesidis

In recent years, the field of data science has emerged as a powerful tool for extracting insights and knowledge from vast and complex datasets. One of the most promising areas in which data science is making significant strides is in the realm of healthcare and medicine.



The Data Science Division at the Center for Molecular Fingerprinting is at the forefront of this innovation, exploring the extensive possibilities of utilizing photonic data in medical diagnostics, personalized health monitoring, and life sciences. Data scientists at CMF are engaged in a comprehensive exploration of how photonic data can revolutionize medical practices. This involves investigating various aspects, including study design, data processing techniques, and the integration of machine learning methods with medical statistical theory. Properly combining these aspects results in powerful data-science pipelines. CMF's strength lies in the direct application of these pipelines in real-world clinical studies. This approach ensures that the methods developed are not just theoretical constructs but are practical and effective in addressing actual medical challenges. These clinical studies provide a valuable testing ground, allowing the division to refine their approaches and enhance the accuracy of their techniques.

The Data Science Division's endeavors, however, extend beyond direct practical applications and into the realm of more fundamental problems in medical decision-making. By leveraging concepts from information theory, decision theory, and statistical physics, the team aims to quantitatively assess the medically relevant information contained within different types of health datasets. This pursuit is driven by the quest to understand the true value and significance of data from various sources in informing medical decisions. Each individual's health profile is unique, and accurately defining their health status requires a comprehensive understanding of all data available. By cross comparing the efficiency of different data sources in defining an individual's health status, CMF seeks to identify the most effective approaches for tailoring medical decision-making to each person's specific needs.

The data-driven approach combined with large clinical studies, conducted by CMF, enables the team to address the challenge of personalized medicine.

Harnessing the power of machine learning

CMF's data scientists aim to transform healthcare through predictive models driven by insights from observational studies and precise infrared measurements of human blood. These models hold immense potential for disease detection and health monitoring, utilizing datasets brimming with valuable information.[2] Powered by machine learning, these models decipher intricate data patterns that are imperceptible to the human eye. However, transitioning from controlled laboratory environments to clinical applications poses challenges. Variations in data distribution and patient populations demand adaptability.

Domain adaptation, a subfield of machine learning, is key.[3] Its purpose is to tailor models from controlled labs to real clinical settings, ensuring robustness. Transfer learning enriches this adaptation, leveraging knowledge to fine-tune models for specific clinical tasks. At the same time, another technique, known as active machine learning, can refine model performance by intelligently selecting informative samples for labelling. This method accelerates development and keeps models current in the dynamic clinical landscape. By employing such techniques, data scientists at CMF strive to bridge the gap between controlled environments and real-world complexities and give momentum to the transformation of healthcare.

---

[2] *Huber, M. et al. Elife 10 2021, "Infrared molecular fingerprinting of blood-based liquid biopsies for the detection of cancer." , Huber, M., et al. Nature communications 2021, 12.1, "Stability of person-specific blood-based infrared molecular fingerprints opens up prospects for health monitoring."*

[3] *Quinonero-Candela, J., et al., Mit Press 2008, eds. Dataset shift in machine*

Unveiling hidden complexities

In the realm of medical diagnostics, the fusion of molecular fingerprinting through vibrational spectroscopy and machine learning presents unprecedented potential. Recent investigations showcase the promise of these approaches, envisioning new diagnostic tests.[4] Yet, the focus has primarily been on medical test performance in controlled settings, overlooking the influence of confounding factors beyond disease status. Current research, based on collaborations between molecular fingerprinting and machine learning, offers only a glimpse into the potential of utilizing photonic data in healthcare. Various complexities that originate from human physiology and environmental factors, can arise and introduce confounding variables. Despite the promise of these techniques, current research delves only sparsely into the impact of factors beyond disease status, ranging from lifestyle to experimental biases. Data scientists at CMF and researchers at Ludwig-Maximilians-Universität München (LMU) are seeking to quantify these impacts by employing rigorous statistical evaluations and investigating factors beyond disease status that affect the performance of potential medical tests based on infrared molecular fingerprinting.[5] The goal is to understand both the benefits and limitations of these tests, scrutinizing their performance in complex real-world scenarios. As this field evolves, its significance is becoming clearer.

Navigating medical decision-making with information theory

In the evolving landscape of healthcare, the ability to foresee transitions from wellness to disease as early as possible takes center stage. Information theory, traditionally applied to communication and data compression, has found an unexpected application in medicine. It offers a unique perspective into analyzing and understanding medical datasets by revealing patterns and relationships not evident through conventional methods. At its core lies the pursuit of precise wellness assessment. Data scientists at CMF and LMU are striving to unravel the essence of wellness, scrutinizing data intricacies to identify markers that distinguish optimal health from early signs of deviation. Guided by probability and information theory, they delve into the mechanics of these transitions, seeking subtle signs that foretell shifts. This goes beyond disease identification, aiming to create a holistic framework for comparing different sources of medical data and assessing their utility in informing early interventions. The mathematical discipline of information theory allows for the quantification of the medically relevant information captured by various medical data panels. In particular, the concept of entropy, which lies at the core of the theory, can be used to assess information content.[6] This approach goes beyond theory; it is grounded in practicality. CMF data scientists are developing a methodology based on the concept of entropy for assessing the amount of medical information from different sources, such as clinical laboratory panels, proteomics profiles and infrared fingerprints. This extracted entropy is directly related to the number of different phenotypes that can be distinguished given a specific health-related data panel. Using this methodology, the utility of different data panels can be evaluated and cross-compared based on their capacity to precisely defining the health state of individuals and potentially capturing signals of transitions from wellness to disease. At the time of writing, medical information from about 10,000 healthy individuals (in a longitudinal setting), has been collected and the quantification of information content of different panels, such as clinical chemistry and molecular fingerprints,is under way.

Molecular fingerprinting combined with machine learning holds immense potential, but addressing confounding factors is pivotal to fully realizing its innovative capabilities in medical diagnostics.

[4] Huber, M., et al. Elife 10 2021, "Infrared molecular fingerprinting of blood-based liquid biopsies for the detection of cancer.", Kepesidis, K. V., et al. BMC cancer 21 2021, "Breast-cancer detection using blood-based infrared molecular fingerprints."

[5] Gigou, L. (2023) Effects of covariates in infrared molecular fingerprinting for cancer detection.

[6] Shannon, C. E. The Bell system technical journal 1948, 27.3, "A mathematical theory of communication."

Unleashing the power of AI in in-silico studies

A category of artificial intelligence algorithms, generative models (GMs) are renowned for crafting authentic synthetic data.[7] They are designed to perform a very particular type of task: to generate artificial but realistic data based on large sets of real observations. Such algorithms are revolutionizing various fields of research, ranging from pharmaceutics to theoretical physics, by producing artificial events that mimic real-world complexities. Such artificial events preserve the statistical and physical characteristics of the original ones. This is achieved by trying to approximate the underlying statistical process hidden behind the real events. In pharmaceutical research, clinical trials are pivotal but resource intensive. In this field, researchers have experimented with GMs to be used in *in-silico* clinical trials[8], which, in many cases, allow the dramatic acceleration in the development of new drugs and medical devices while also significantly cutting R&D costs. Traditional clinical studies demand time and resources, but *in-silico* studies compress timelines and accelerate discovery. Financially, *in-silico* studies powered by generative models alleviate burdens, reducing material, labor and logistics costs. This approach democratizes clinical research, fostering innovation accessible to diverse researchers and organizations. At CMF and LMU, data scientists are developing GMs for the creation of artificial patients that possess well-defined demographic characteristics, such as specific age, gender and body-mass index (BMI), as well as specific disease status (lung cancer, diabetes etc.). These simulations can replicate different types of clinical studies, such as longitudinal cohort studies or case-control, allowing researchers to test interventions and hypotheses in a wide range of settings in a risk-free manner.

Medical diagnostics through innovative probabilistic models

The pursuit of optimal medical diagnostics continues to drive innovation. Developing an unbiased and efficient diagnostic approach demands a thorough, probabilistic assessment of risk factors and indicators, including precise molecular profiles. Similar challenges have engaged physicists for years, as they dissect macroscopic states of diverse physical systems through the scrutiny of microscopic components and their interactions. These insightful lessons could potentially find an application in the medical domain.[9] Collaborative efforts between the LMU and CMF aim to develop models based on the convergence of methods from statistical physics, machine learning, and inference algorithms. This convergence could enhance the existing methodologies for medical diagnosis. A potentially transformative project currently starting at CMF aims to do exactly this: redefining the understanding of diseases and related diagnostic procedures through probabilistic models that delve into disease complexities beyond binary classifications. These models can offer nuanced insights into aspects of disease and reveal optimal ways of applying diagnostic testing.

---

[7] Goodfellow I., et al. MIT press, 2016, Deep learning.

[8] Wang, Z., et al. (2022) arXiv preprint arXiv:2209.09023 "Artificial Intelligence for In Silico Clinical Trials: A Review."

[9] Ramezanpour, A., et al. Diagnostics 2020, 10.11, "Statistical Physics for Medical Diagnostics: Learning, Inference, and Optimization Algorithms.")

The road ahead: a future shaped by knowledge

The Data Science Division at CMF stands at the forefront of an exhilarating journey that brings together new technologies, data science, AI and healthcare. Our research aspires to drive a wave of innovation that promises to redefine diagnostics, treatment, and health monitoring. As we look ahead, the collaborative efforts of CMF, LMU, as well as partnering research institutions have the potential to revolutionize the way we approach healthcare, offering a future where precision, early intervention, and accessibility converge for the benefit of patients worldwide.

Women in science

– interview with Dr. Mihaela Žigman

> *...moving mountains requires a team effort!*

"I very much hope that science is a profession where our capacities can overcome questions of gender and where scientists are professionally acknowledged commensurate with their intellectual capacities." - said Dr. Mihaela Žigman, the Research Director of CMF. But in reality, only 33% of researchers globally are women. Women are actively pursuing bachelor's and master's degrees and even outnumber men at these levels, as they represent 53% of graduates. Still, their numbers drop off at the PhD level to 43%, and the discrepancy widens at the researcher level. The high proportion of women in tertiary education does not necessarily translate into a more significant presence in research, according to the UNESCO Science Report: towards 2030. Many countries recognize that a more even gender balance and diversity in science and research would increase their competitiveness. While states and governments are working on this, raising awareness at the societal level is also essential. A great example is the International Day of Women and Girls in Science, celebrated every year on February 11 to help recognize the critical role women and girls play in science and technology and support gender equality.

Why is the International Day of Women and Girls in Science important?

Celebrating this day also entails recognizing the importance of diversity in thought and collaboration, essential elements for meaningful scientific progress. As we continue to push the boundaries of science, let us reflect on the fact that women have not only participated in scientific endeavors throughout history but have also been instrumental in shaping our understanding of the world and the way we practice science today. There is no doubt that countless examples of invaluable contributions by female scientists exist.

In the present day, my concern has shifted from the inclusion of female scientists in industrialized nations to a broader problem. I worry that the opportunity to learn, study, and pursue an intellectual career is often a socioeconomic luxury that many people can't afford. In some countries, this challenge is even more pronounced, particularly for girls.

What inspired you to become a scientist?

I think the inherent wish to better understand the phenomena around us to comprehend how nature functions has always intrigued me.

Was your journey challenging in a field where there are fewer women?

In my work environment, particularly in the Natural Sciences and Life Sciences, I would say that I had equal opportunities as a woman to advance in my career. I've worked as a scientist in both female-dominated and male-dominated research laboratories at various European and US institutes. Overall, I never felt limited or challenged due to my gender. I strongly believe that science should be a profession where our capabilities can transcend questions of gender and where recognition should be based on intellectual merit.

What was the best advice that you received early on in your career?

It is essential to identify a topic that personally fascinates you so that you can thrive as an individual when studying it. Whatever that phenomenon, problem, or theme is, it will most likely be of secondary importance, but you need to get personal satisfaction from working on the issue you have chosen.

What would be your advice to girls who want to become scientists?

If you have a deep interest in a subject, keep exploring it further. Dive deeper! Learn how to ask simple, empirically answerable questions such as "Why is it like that? How does it work, and what is its significance?" This inquisitive attitude, innate in every child, should never cease. Pursue your curiosity, and surround yourself with people who are cleverer than you, especially those who don't flaunt it. That's when you begin developing the means to find the answers to your questions.

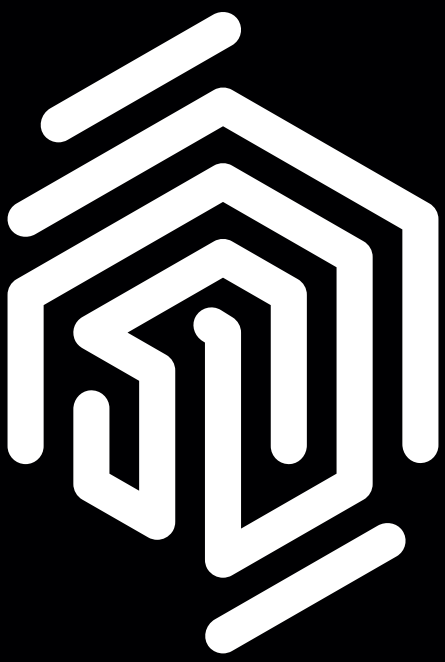
The winner of the 2023 Nobel Prize in Physics: Prof. Dr. Ferenc Krausz

The 2023 Nobel Prize in Physics has been awarded to Ferenc Krausz, Pierre Agostini, and Anne L'Huillier "for experimental methods that generate attosecond pulses of light for the study of electro dynamics in matter", The Royal Swedish Academy of Science announced on October 3, 2023.

The three Nobel Laureates in Physics 2023 are being recognized for their experiments, which have given humanity new tools for exploring the world of electrons inside atoms and molecules. Ferenc Krausz, Pierre Agostini, and Anne L'Huillier have demonstrated a way to create extremely short pulses of light that can be used to measure the rapid processes in which electrons move or change energy.

Ferenc Krausz pushes the boundaries with a relentless drive for progress. His vision is championed by the entire team at the Center for Molecular Fingerprinting. Together, their mission is to shift the boundaries of disease detection, paving the way for personalized preventive medicine. Each team member plays a vital role in shaping the future of healthcare, and they are dedicated to transforming their collective vision into reality.

Photo credit: Árpád Földházi / Mandiner



LASER PHYSICS
FOR HEALTH

HUNGARIAN NATIONAL
LABORATORY

www.cmf.hu