# Predictive Medicine by Cytomics

***former:* *Cell Biochemistry Group***
- ***Max-Planck-Institut für Biochemie*, *Martinsried***

# Individualized prediction of disease progression and outcome

*(Evidence Based Medicine at the Cellular Level)*

- ***< Cell Biochemistry*** **(PDF)**
  - ***= external links***

## 1. Aims and Potential

**1.1** Pharmaceuticals are typically developped according to *best group (cohort) efficiency*. Once approved they are applied to similar groups of patients. Some patients may, however, not benefit from a presently optimal therapy and are potentially harmed by unwanted therapeutic side effects (adverse drug reactions (ADRs)) despite the improved *prognosis* (=*group future*) of the entire patient group. This is suboptimal. Accurate *predictions* for the reactivity of the *individual patient* in such groups prior to therapy onset constitute therefore a *primordial goal* of predictive medicine by cytomics. Individualized prediction of disease progression (disease course prediction, outcome prediction) will improve *overall therapeutic efficiency*, better comply with the "*primum nil nocere*" principle in medicine and meet the *central patient interest* to be *cured* of disease by an *individually optimized therapy*.

**1.2** *Predictive medicine by cytomics* (*molecular cell system analysis*) (**fig.1**) aims at > 95% or higher accuracies for therapy related disease course or outcome predictions in individual patients by differential *data pattern classification* (*predictive differentials, predictive differential classification*) of molecular cell phenotypes or other molecular measurements in patients. Cells constitute the *elementary function units* of cell systems (*cytomes*), organs and organisms. *Diseases* are caused by molecular changes in cells. This means for the detection of early disease processes: ***cells know it mostly first***. Cytometry measurements can detect such altered *molecular cell phenotypes* resulting from *genotype* and *exposure* influences. In case disease inducing cells are not accessible, *disease specific molecular patterns* of immune *indicator cells* like cellular or humoral responses of lympho-/monocytes or granulocyte activation in blood or other body fluids can be probed instead.

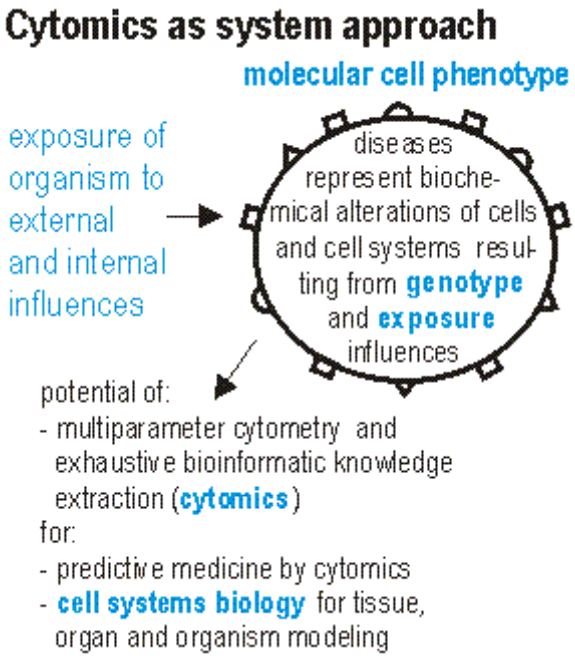


**fig.1** System cytometry and cell systems biology

*Similar* diseases may result from high genotypic susceptibility and low exposure to external influences or alternatively from low genotypic susceptibility at high exposure. The *high* genotypic diversity in man at a comparatively *low* number of possible diseases emphasizes the *potential* of *molecular cell phenotypes* as *diagnostic*, *therapy guiding* and *outcome prediction* indicators in individual patients. Instead of trying to cure patients according to *individual genotype*, it may be more promising to therapeutically address disease specific *molecuar cell phenotypes* thus considerably reducing the number of potential therapies.

It is in any case *questionable* whether future *disease occurrence* can be predicted from *genotypes alone*, at a time where the *future exposure history* of an individual is still unkown. Exposure to external influences is an important disease promoter as evidenced for example by the non uniform occurrence of *morbus Parkinson* in identical twins (**ref.1,2**). Altered *molecular cell phenotypes* may, even in this case, provide *earlier* information about *future disease occurrence* than genotype determination alone.

**1.3** Altered molecular cell phenotypes are determined as *differential* classification masks by iterative selection of the most discriminatory triple matrix patterns between diseased and healthy patients. The optimization process provides *disease and patient classification masks* (rightmost table columns) (*hotspot heat masks*). They represent direct or indirect *molecular equivalents* of *disease processes*. Classification masks can be established for diseased or disease associated cells like inflammatory immune cells. Either patterns may *vary* to a certain degree from patient to patient due to different combinations of genotype and exposure influences. This does, however, not influence the accuracy of the robust classification process. The *individually optimal therapy* (*individualized medicine, personalized medicine*) can be selected by data pattern classification of patient groups *stratified* for example according to Kaplan-Meier. The presented concept of personalized medicine concerns the care of *diseased patients* or of persons during *disease development*. It does **not** aim at the prediction of future disease occurrence from the person's individual genotype (*transparent patient, vitreous man*). The concept has a *wider application potential* than the *pharmacogenomics* or *predictive medicine by genomics* approaches of personalized medicine. The *algorithmically* determined data patterns can be standardized. **No** statistical or correlation (dendrogram) analyses are used for the classification process.

**1.4** Patients with a prediction for "*disease aggravation*" may convert under therapy within some time to "*non-complication*" patients such as e.g. in intensive care medicine. The early detection of disease aggravation or amelioration provides a *lead time* for preventive therapy onset or for therapy reduction (preventive medicine).

**1.5** Therapeutic *lead time* may increase overall therapeutic efficiency by the prevention or reduction of disease induced irreversible tissue damage or of unwanted therapeutic side effects. It may also permit to identify risk patients *prior* to disease declaration like in asthma, rheumatic diseases or diabetes. This may help to *delay* disease outbreak and *reduce* complication rates as an important practical consequence.

**1.6** Accuracy levels for individualized predictions of disease progression can be increased in principle from presently around 95% to 99% or higher upon merging the most informative parameters from different studies into the disease classification masks (*"disease signatures"*). The *knowledge extraction* by data pattern classification is independent of mathematical assumptions concerning the value distribution of parameters and the optimal classification is obtained unsupervised that is in an automated way with high certainty for the selection of the correct data pattern. The classification is also comparatively robust against the misclassification of random statistical aberrations as true aberrations.

**1.7** The two-step research strategy consists of **i)** *hypothesis-driven* (*deductive approach*) determination of experimental *molecular cell phenotype* parameters of diseased and healthy individuals, followed by **ii)** *hypothesis-free* differential data pattern classification (analysis, mining) for all investigated cells in their full *heterogeneity*.
The use of healthy patients as *reference groups* permits the elaboration of standardized classifiers (*periodic system of cells*) by the combined reclassification of the most discriminatory parameters of several experimental approaches, performed under different hypotheses (*inductive approach*). Non cellular molecular parameters for example from blood serum, urine or liquor may be additionally included in the analysis. Data patterns with more and more discriminatory efficiency are obtained in this way (*autofocusing*). This may permit to identify new disease associated *molecular hotspots*, being presently *inaccessible to hypothesis development* due to the lack of preexisting knowledge ("*observing molecular medicine"*).
This *concept* and *data driven* molecular *top-down* approach is initially comparatively *independent* of prior knowledge about the ultimate molecular causes of disease. In particular there is *no need* to first analyze the molecular effects of *hypothesis driven* systematic perturbations of cellular *model systems* as they are frequently used to acquire knowledge about disease affected molecular pathways. Subsequently these pathways are investigated in detail by the *bottom-up* concept of *systems biology* (*system biology*). Molecular *cytome exploration*, in contrast, analyzes *differential molecular disease patterns* in patient cells. The detour of investigating molecular pathways in potentially *unsuitable cellular model systems* is avoided, information on *therapy dependent future disease development* in individual patients is obtained with the potential to *simplify* investigations on disease mechanisms by favoring the development of new hypotheses.

**1.8** Once a certain molecular knowledge has been assembled, disease inducing molecular pathways can be explored by a *retrograde molecular analysis strategy* (*molecular reverse engineering*) of molecular cell phenotype differentials at the *cell system level*. The pathways can be mathematically modeled (*biomedical cell systems biology*) to further increase the predictive capacity. It is likely that new target molecules and lead structures for *drug discovery* will be detected by *hypothesis-free* data pattern classification due to its capacity to address *unknown molecular knowledge spaces*. In this sense *cytomics* represents an entry to *biomedical cell systems biology*.

**1.9** The described classification concept *concentrates* the differentially most informative molecular cell parameters within *specific disease classification masks* containing typically between 5 and 30 parameters. It does *not* advocate for the determination of ever increasing parameter sets generating frequently *interpretation* difficulties at the individual patient level. An initial goal of this effort is to build a system of standardized, inter hospital exchangeable and individually predictive data classifiers for patients, possibly within the framework of a human cytome project.
The *potential* of the concept consists in its general applicability to various areas of clinical or ambulant medicine as illustrated below (**chapter 2**) by collaborative projects with individual hospitals and institutions as well as within the framework of the European Working Group on Clinical Cell Analysis (• EWGCCA) in the context of clinical cytomics. The apparent challenge is to advance this effort to the patient level in a multistep effort of scientists, clinicians and industry as proposed in the context of the ***human cytome project*** (**PPT**, **ref181**, **ref175**, **ref170**, **concepts**, **definitions**, **cytomics references**) or in the establishment of a *periodic system* of cells with stem cells or other cell compartments as reference. Despite ressemblance in name, this concept differs significantly from the earlier concept for a • ***plant periodic cell system***.

A human cytome project may induce the elaboration of a **standardized disease classification system**.
The number of human diseases is in the *hundreds* or *thousands*, that is significantly *inferior* to the several billions of individuals on this planet. Many diseases manifest in multitudes of *ethnically* and *genetically* different patients with different disease histories and exposure to environmental influences during theit lifetime. This leads to *heterogeneities* in therapy response like in rheumatoid diseases or malignancies. *Clinical medicine* tries to cope with this situation by pretherapeutic *patient stratification* to determine as good as possible the *most susceptible* patients. A *standardized disease classification system* based on optimized molecular data patterns has the potential to define diagnostic entities *more precisely*, including

therapy outcome prediction for individual patients.

### 2. Individualized prediction of patient disease progression (Medical Cytomics, Clinical Cytomics)

- **pretherapeutic identification of high risk AML patients**
- **pretherapeutic identification of high risk DLBCL patients**
- **identification of high riskcolorectal cancer patients**
- **disease activity and prediction of therapeutic efficiency in SLE patients**
- **salicylate & hyperthermia potentially inhibit disease progress during influenza or COVID19 incubation period**
- **outcome prediction in sepsis patients**
  - **ARTE TV** (no more on-line)
- **preoperative identification of risk patients for postoperative effusion and edema (POEE) in children cardiac surgery**
  - **3sat Nano TV** (no more on-line)
- **prognosis of melanoma patients**
- **risk assessment for overtraining syndrome in competition cyclists**
- **risk assessment for myocardial infarction**
- **classification of leukemia and lymphoma**
- **classification of immunophenotypes and clinical chemistry parameters in juvenile asthma**
- **staging of HIV patients from immunophenotypes**

### 3. Non Medical Data Classification

- **microplankton classification in ocean waters**

### 4. References

**1.** *CM Tanner, R Ottman, SM Goldman, J Ellenberg, P Chan, R Mayeux JW Langstorf.* Parkinson disease in twins: an etiologic study. JAMA (1999) **281**:341-346.
**2.** *K Wirdefeld, M Gatz, ChA Reynolds, CA Prescott, NL Pederson.* Heritability of Parkinson disease in Swedish twins: a longitudinal study. Neurobiol Aging **32**(10):1923:e1-1923.e8.doi:10.1016/j.neurobiolaging.2011.02.017.

### 5. Timeline: Evolution of Concept

---

- *< Cell Biochemistry*

---

## *Off-line Internet, a time saver !*

**Download** the ZIP file: **classim1.zip** containing all Cell Biochemistry pages for example into directory: **d:\classimed**, unzip into the same directory, enter the address: **file:///d:/classimed/cellbio.html** into the URL field of the Internet browser to directly access text & figures on your harddisk free of network delays (further information).

---

Impressum © **2020 G.Valet** Mail

*Internet: https://www.classimed.de*
*last update: Mar 30,2020*

*first display: Jan 10,1995*

---