Images in medicine is not only X-rays and PET/MR/CT...

Cervix cancer screening

- In the 40-ies Papanicolaou developed a method for visual screening of specially stained cell specimens for early detection of cervical cancer, the ‘PAP-smear’.
- If the cellular changes are detected early the cancer can be cured.

**Normal**

Abnormal cells have nuclei that are:
- larger
- more irregularly shaped
- coarser in texture
- relatively smaller cytoplasms

All women in developed countries are tested every 3-5 years from age ~20-65. In the US this means ~150 million samples per year (2-3 million abnormal).

Can the analysis be automated?

The Cytoanalyzer, 1956, US

an electronic device that screens samples of cervix smears of suspected malignancies.

Automated cytology - early systems

Around 1980 several automated pre-screening devices based on algorithms developed during the 1970-ies and implemented in special digital hardware arrived:
- CYBEST 4 in Japan
- BioPEPR in Holland
- FAZYTAN in Germany
- CERVIFIP in Great Britain
- LEYTAS in Holland
- DIASCANNER in Sweden
(by Ewert Bengtsson, IMTEC)
Automated cytology - early systems

Around 1980 several automated pre-screening devices based on algorithms developed during the 1970-ies and implemented in special digital hardware arrived
- CYBEST 4 in Japan
- BioPEPR in Holland
- FAZYTAN in Germany
- CERVIFIP in Great Britain
- LEYTAS in Holland
- DIASCANNER in Sweden (by Ewert Bengtsson, IMTEC)

All the systems were field tested but none became a commercial product.

Finally commercial screeners

In the end of the 90-ies the first successful commercial screening systems appeared.
- The BD FocalPoint™ Slide Profiler Intelligent Pap Imaging directs attention to slides most likely to contain abnormality and can eliminate clearly normal slides from visual inspection.
- Hologic ThinPrep™ promotes Dual Review, all slides are screened by machine and visually.

Before the Pap smear was introduced into clinical practice, carcinoma of the cervix was the leading cause of cancer related deaths among American women. Now it is down to being number eight on the list: the most successful cancer detection strategy ever developed.

The systems are too expensive for use in poor countries, and cervical cancer is still a leading cause of cancer deaths in women where Pap smear screening is not widely available.

Why did the development take so long?

Difficulties with displaying digital images, data storage, computer resources, lack of robust methods, and acceptance in the clinic.

Before early 1970-ies no graphical displays were available. Images were displayed on paper by overprinting characters on text printers.

‘EXCELL interactive cell image analysis’, Bengtsson et al, Uppsala 1984:
- Had one of the first 1024x1280 full colour displays, custom built, each display costing around 10 000 Euros.

Image based High Throughput Screening (HTS)

- A well established approach to test the effect of a large number of chemical (or genetic) perturbants to discover new treatments or help us understand biomedical mechanisms. HTS, High Content Analysis, refers to the number of different parameters (shape, intensity, texture, localisation etc) that can be extracted from image data.

1. Add cells to each well in a 384-well plate.
2. Check or DNA treatment by microscopy + image analysis.
3. Repeat.
4. Find ‘hits’ and explore potential impact on human health.

Practical approach to HTS on cells

‘Measure first, ask questions later’

10^5 images, 10^4 cells in each: Total of 10^6 cells/experiment.

Delineate each cell and measure a large number (~500) features (size, shape etc) per cell (~“cytoprint”). Then use these measurements to find ‘hits’—orals with a desired change in appearance.

Robots add potential drugs from chemical libraries to each well of the multi-well plate, wells are imaged, and analyzed to search for desired effects on the cell level.

384 wells/plate
Example: Breast cancer
A search for inhibitors of heregulin, which stimulates the ErbB2 pathway and causes cells to be migratory (=bad)

What can we do if we can collect 100,000+ images per day?
• Test a million chemicals to see which impede cancer growth
• Test a million chemicals to see which prevent neural degeneration
• Test every gene in the genome to see which are required for proper metabolic function
• ...essentially conduct any microscopy experiment, for the whole genome or for a million chemicals

Experimental design
– What to think about when approaching a new problem
– Hypothesis: why is it important?
– Sample preparation
– Quality control
– Evaluation/Benchmarking of automated image analysis approaches

A future scenario:
Can you write a program that automatically counts all the dots in these images?

A better future scenario:
Sure. How many dots are there?
Hypothesis

Wikipedia:
A hypothesis (from Ancient Greek ὑποθέσις, from Greek ὑποθέτω — hypothetical meaning "to put under" or "to suppose,"[1] plural hypotheses) is a proposed explanation for a phenomenon. For a hypothesis to be a scientific hypothesis, the scientific method requires that one can test it. Scientists generally base scientific hypotheses on previous observations that cannot satisfactorily be explained with the available scientific theories.

A hypothesis can also be thought of as an 'educated guess' for which one wants to find evidence.

By formulating a hypothesis, pinpointing the goal of an experiment becomes much easier, and will facilitate experimental design.

Scenarios of different hypothesis

Hypothesis: The number of dots changes at a given treatment...
- Do you have positive (treated) and negative (untreated) controls?
- Do you have more than one positive and negative control image so that I can measure the variance within the controls?
- Hypothesis: Dots appear when adding X.
  - Do you have any images without dots (only background)?
  - Hypothesis: Dots cluster and increase in size when adding X.
  - Additional measurements such as shape, size and intensity may better quantify the changes?
- Hypothesis: Image-based measurements of Y (hypothesis above) is more powerful than the standard approach 2.
- How can we benchmark against 2?

A clear hypothesis will be helpful for experimental design

- What is the trade-off when it comes to improvements/consistency of sample preparation and image quality
- Higher resolution vs number of spots per field
- Using auto-focusing
- Imaging all samples using the same illumination source
- Avoiding glare and shadows etc
- Is this collection of images representative of the images you want to analyze?
- What do the worst images look like?

Sample preparation and imaging

A digital image is characterized by
- geometric properties and resolutions (x,y,z)
- densitometric and spectral properties and resolutions (λ)
- temporal properties and resolutions (t)

... and the preparation of the imaged sample.

All these different aspects affect the analysis, and choice of sample preparation and imaging sensor are crucial for any experiment.

Data quality control

- If we have 100 000 images, and 10% of them are out of focus or contain debris that skews our measurements, there is a large risk of getting many false hits and also missing hits.
- How can we find these ‘bad’ images?
- What to do once we’ve found them?

Three important means of QC

- Good experimental design
  - Helps to identify systematic errors (especially those linked for example with well position) and determine what normalization should be used to remove/reduce the impact of systematic errors
- The selection of effective positive and negative (chemical/biological) controls
- The development of effective QC metrics to measure the degree of differentiation so that images with inferior data quality can be identified and flagged/excluded.
Quantify blur

- What is a good way to find images where cells are out of focus?
- How can we make the method independent of the number of cells in the image?

Examples of quality measures

- We want a measurement that is invariant to cell count and has its extremum when the image is well focused.
- Mean/STD or 'S/N' – Calculated automatically by imageprocessing (commercial system for HTS) but very inaccurate.
- Focus score – Normalized intensity variance – Recommended for auto-focusing
- Correlation (locally) – Must specify spatial scale
- Power LogLog Slope (PLLS) – Evaluates the slope of the power spectrum density of the pixel intensities on a log-log scale

Evaluation and benchmarking

- Crucial for methods development
  – Does my proposed approach support the hypothesis?
  – Is this approach 'better' than a previous approach?

Approaches for benchmarking your methods development:

Counts: number of objects in each image is provided, as counted by one or more humans

Foreground and background: a human produces a binary representation of foreground and background. Evaluation: calculate precision and recall, and rank algorithms by F-factor (i.e., the harmonic mean of precision and recall).

Outlines of individual objects: a human outlines each object in the image in order to indicate which pixels belong to which object. Evaluation: consider each pixel that is on the boundary of the object found by the algorithm and sum up the distances to the closest pixels in the human outlines.

Biological labels: positive and negative control treatment.

Visual annotations

Commercial software vs Free and open-source

Many microscope systems come with built-in software for quantitative measurements – use these if they can do what you want to do!

But...

"The most fundamental element [of open-source software] is the openness: if you can't see the code of a piece of commercial software, then you cannot say what the software really does, and this is not scientific."

– Andrew C. Gates (supporter of Fiji)

"Reproducible research": with an open-source solution, you can provide your analysis pipeline as part of the supplementary material of your published paper (along with data).
Biological imaging software tools

A large number of software to enable
• acquisition
• management
• analysis
• visualization


About

CellProfiler

cell image analysis software

CellProfiler Analyst

data exploration software

• Available for free from www.cellprofiler.org, developed at the Broad Institute of Harvard and MIT, supported by grants from NIH and other.
• Designed for HTLCA, but 70% of users do low medium throughput (~100 images)
• CF downloaded > 25 times per calendar day and launched more than 200 times per calendar day (excluding branches from the Broad Institute)
• Open source (written in Python, communicates with images)
• Software available for Windows, Mac and Linux.
• To get started, download example images and pipelines from www.cellprofiler.org and take a look at one of the video-tutorials.
• To get help, contact the user forum at http://cellprofiler.org/forum

Example

Control experiment: We have a full plate with 96 wells containing human adenocarcinoma cells, a positive control drug (a drug known to promote cell proliferation) and a negative control drug (a drug known to inhibit cell proliferation). We know that the positive control drug causes the cells to divide, and the negative control causes them to die. We want to optimize the cell count to separate positive and negative controls as well as possible (measurement 7).

A modern example

2012: The Human Protein Atlas Project

The Human Proteome Atlas: ‘Google Maps’ for proteins, from the subcellular scale, to the tissue- and organ-level
Why is localization important?

- Localization is closely related to function, and understanding function can help us understand mechanisms of action.
- The presence or absence of the ‘wrong’ protein in a given location can help diagnose disease.
- Understanding the mechanism of action facilitates design of drugs that cure disease.

“All current drugs with a known mode of action act through 324 distinct molecular drug targets. Of these, 266 are human genome-derived proteins, and the remainder are bacterial, viral, fungal or other pathogenic organism targets.”


Why is localization challenging?

- Large scale; the human genome is believed to contain approximately 20000 protein coding genes. Each of these protein coding genes can however give rise to a much larger number (millions) of different proteins. This calls for automation.
- Large variability in abundance. This calls for quantification.
- Large variability in detectability; two proteins may not be possible to detect using the same protocol. This calls for optimization.
- The human proteome is largely unexplored. This calls for validation.

Annotations done manually by the Mumbai HA team

Dr. Sanjay Navani

High Quality Pathology Annotations
Fruitful scientific collaborations
Where is the automated analysis of histological samples by digital image processing?

Complex structures, variability in sample preparation and variations in staining intensity makes automation tricky.

What can be done to make analysis of tissue easier?

More specific molecular probes.
- Signals that are easier to quantify.

Take home message

Knowledge about the image formation, possibilities and limitations, can greatly improve the scientific value of an experiment.

A better understanding of digital image processing, possibilities and limitations, can greatly improve the scientific value of an experiment.

A better communication between experts of from different fields can greatly improve the scientific value of an experiment.

"With the aid of informatics, microscopy is in the midst of a crucial evolution into a more quantitative and powerful technique."

Daniel Evans, editor of Nature Methods, June 2012