Advances in Instrument Intelligence

Using AI to Achieve Digital Transformation

WILEY

Agilent

Contents

3 Introduction
4 The "Industrial" Revolution in Biomedical Research
7 Artificial Intelligence (AI) Transforming Laboratories
9 Depicting the Most Recent Intelligent Workflow Tool to Maximize Analytical Throughput: iReflex

Imprint
© Wiley-VCH GmbH
Boschstr. 12,
69469 Weinheim, Germany
Email: info@wiley-vch.de
Editor-in-Chief: Dr. Christina Poggel

Advances in Instrument Intelligence

Introduction

The laboratory workflow is changing rapidly to keep up with today's fast-paced world. Thus, it is necessary that products and people also adapt, not only to increase productivity but also to eliminate anomalies or inconsistencies induced by human error and variations in processes.

The main challenges include the requirement of increasingly sophisticated quality control to restrain errors, combined with advanced analytics to obtain meaningful information or reliable diagnostics. Data reproducibility is also critical since failures in reproducing and validating results threaten the integrity and reputation of scientific research. Automation must merge with human expertise.

Instrument intelligence can help in many ways. For example, depending on the nature of data and the processes of obtaining and storing data. On key resource is the implementation of a tool based on machine learning and statistical data analysis that can be used to highlight unusual values or suspicious sequences of values, identify inconsistencies over different data sources and data modalities, and enable reducing the number of tests by automatically filling in the missing values or pointing out where more data should be collected.

This eBook presents an overview of instrument intelligence solutions and shows how several lab sectors can be modified to have faster and error-free production. This can be especially achieved with digitalization and digital transformation.

The content of this eBook consists of a summary of Wiley's book "Digital transformation of the lab", which discusses opportunities, needs, and challenges related to the role of artificial intelligence (AI) in transforming laboratories, a summary of an article on the "industrial" revolution in biomedical research, and an infographic on Agilent's iReflex intelligent workflow tool.

The "Industrial" Revolution in Biomedical Research

Data explosion and reproducibility crisis drive changes in lab workflows

Several factors are driving profound changes in the way life science laboratories organize their workflow, whether in medical diagnostics or basic research. A common cause of these changes is the proliferation of the amount of data generated combined with a rapid decline of the costs of data production. This requires increasingly sophisticated quality control to restrain errors combined with advanced analytics to obtain meaningful information or reliable diagnostics. The other challenge that imposes changes on laboratory workflow is reproducibility, which is critical since failures in reproducing and validating important results threaten the integrity and reputation of biomedical research.

Deep Learning and Automation

The rise of deep learning (DL) is increasingly demanding for production of accurate data sets to avoid biases and false conclusions. This, in turn, requires greater use of automation to eliminate anomalies or inconsistencies induced by human error and variations in processes. For example, batch effects – which undermine DL work – get much worse without automation. Algorithms that can mitigate batch effects often use some form of Bayesian inference that compares the results of experiments under varying conditions to filter out inconsistencies associated with the process. However, these algorithms can also eliminate biologically significant variations. Growing automation avoids this by eliminating many causes of batch effects in the first place and, thus, prepares the ground for robust application of DL.

Link Different Expertise

Nonetheless, the challenges of workflow do not stop at automation and data analysis. The greater problem lies in reconciling automation with human expertise, particularly in diagnostic laboratories. The need for new techniques to help clinicians exploit computational power effectively has become more urgent with the advent of whole genome sequencing (WGS) and whole exome sequencing (WES). To overcome the diagnostic bottleneck, expanded carrier screening panels capable of identifying hundreds of mutations associated with rare genetic diseases have been developed. These panels can find mutations that would otherwise not be detected, and this is where the integration between machines and humans becomes important.

Changes in Workflow and Labor

The way laboratories work will change radically, as demonstrates a recent initiative by Agilent Technologies, a specialist developer of analytic tools and software for lab workflow, created in 1999 as a spinout from Hewlett Packard. Results of their first Pharma Lab Leaders Survey were published in June 2019, based on responses from 650 lab managers, directors and supervisors from China, Germany, India, South Korea, Switzerland, Austria, and the USA, in a study run by Frost & Sullivan. The results have already led to changes in lab equipment, according to Darlene Solomon, Agilent's Senior Vice President and Chief Technology Officer. "85% of respondents told us that they were buying more sophisticated instruments with a greater degree of specificity", she said.

Greater ease of use and reduced need for training are being driven not just by pressures of workflow but also changes in the skill levels of lab technicians Solomon added. "We have examples where a seemingly small change, such as having instruments with touch screens, has really helped with ease of use and ease of training," Solomon said. "This is important as the profile of a lab technician in some regions has changed. While some years ago lab technicians were likely to be science graduates with specialist education in mass spectrometry or gas chromatography, nowadays, they are as likely to be generalists with a humanities degree, which means that lab managers want instruments that are easy to use, easy to train, and do not need to be operated by a specialist".

Another major trend is the growing incorporation of AI and especially DL to adapt to either complex or repetitive tasks in a lab. "Modern AI approaches based on deep neural networks are especially promising in the area of digital pathology based on tissue staining and can improve the accuracy and reduce the complexity of pathology slide interpretation for cancer diagnostics and treatment decisions," Solomon said. "For example, AI can increase efficiency and improve the accuracy of time-consuming tasks like counting cells, or of the identification of intricate patterns generated by advanced cell staining that could be confusing for manual interpretation."

Solomon stressed over the need for reliable data to feed DL algorithms and highlighted the importance of consistency across workflow. "The analysis of samples always starts from sample preparation. Good and consistent sample preparation is one large step towards confident and quality results that will generate meaningful understanding on the back end," she explained.

Addressing Governance and Ethical Issues

There is yet another dimension to workflow beyond the scope of the Agilent survey, which is the role of motivation and ethical considerations to ensure consistent and fair results. Adamson Muula, Professor of Public Health and Epidemiology at the University of Malawi highlighted this. "Without systems in place even the best equipment does not function optimally," he explained. "It may be we can also adopt Joseph Mfutso-Bengo's model of LEGS (Leadership, Ethics, Governance and Systems)." Muula was referring to the framework called LEGS developed in Malawi to strengthen health systems, particularly in developing countries where governance and rule of law are relatively weak[1]. According to the authors of the LEGS study, ethics leads to stronger internal social control at various levels, such as procurement, clinical work, and research.

Improving Reproducibility

While consistency and accuracy are required in the clinic, reproducibility has become a big challenge for basic and translational research. Among major causes of failure to reproduce results are publication bias when positive results are favored over negative findings. There is also the phenomenon known as HARKing (Hypothesizing After the Results are Known), where researchers present unexpected results as if they had been hypothesized from the start[2]. While not all researchers agree that HARking is entirely detrimental to scientific progress, there is an element of dishonesty in it, and more importantly, it can introduce a bias, as the hypothesis might be just one of several that could be deduced from the results. One solution is again to change the workflow and pre-register work plans and hypotheses before an experiment is conducted to avoid selective reporting or HARking.

Among those interested researchers is Marcus Munafò, Professor of Biological Psychology at the University of Bristol, UK, who studies neurological pathways associated with alcohol and drug abuse. Munafò became interested in the link between workflow and reproducibility after his PhD when he noticed that many findings were less reliable than they seemed, after conducting systematic review and meta-analysis methods. "I became interested in how our incentive structures and ways of working contribute to this," he explained. "Now I'm interested in thinking about ways in which we can improve our incentive structures, and research culture more generally, to focus on quality at least as much as innovation, novelty, and discovery. Over the last few years, we have moved towards an open research workflow. We began by pre-registering our study protocols and now we routinely archive data, and we are beginning to archive analysis scripts with these, and we post preprints of all our manuscripts. Initially this was for grant funded activity only; now it's for all our activity, including student projects." Munafò added there was further to go in this ongoing process, such as sharing more study materials. "This has several benefits – it allows greater scrutiny of our work and introduces incentives for closer checking internally," he said. The great challenge, according to Munafò, is to instill a long-term view among researchers and clearly articulate the benefits, so that group members buy in to the effort required. This aligns with the message from the LEGS project in Malawi, with the underlying theme that motivation in research needs recalibrating around consistency and transparency led by science and data rather than hope and aspiration. It is almost a challenge to human nature itself.

References

[1] Mfutso-Bengo, J., Kalanga, N., and Mfutso-Bengo, E.M. (2018) Proposing the LEGS framework to complement the WHO building blocks for strengthening health systems: One needs a LEG to run an ethical, resilient system for implementing health rights. Malawi Med. J., 29 (4), 317.

[2] Kerr, N.L. (1998) HARKing: Hypothesizing After the Results are Known. Personal. Soc. Psychol. Rev., 2 (3), 196–217.

[3] Hunter, P. (2020) The "industrial" revolution in biomedical research. EMBO Reports 21, e50003.

Artificial Intelligence (AI) Transforming Laboratories

Digitalization is spreading in our work and daily life and moving from paper to digital is much more than just changing the medium for storing the data. This article discusses opportunities, needs, and challenges related to the role of artificial intelligence (AI) in transforming laboratories.

Digitalization of laboratories is happening in some areas faster than others, depending on the needs and opportunities to make a change from paper to digital. AI can serve as a catalyzer in the process, offering a whole range of additional services including sanity checking, outlier detection, data fusion and other methods in the data preprocessing phase, different methods for data analytics and modeling, monitoring of data consumption and other dynamic processes in the laboratory, and decision support to domain experts **(Figure 1).**

Data Preprocessing and Data Analytics

AI methods can help in data preprocessing in different ways, depending on the nature of data, the processes of obtaining and storing data, or some other data properties. For instance, we can have a tool based on machine learning and statistical data analysis to highlight unusual values or suspicious sequences of values, identify inconsistency over different data sources and data modalities, and enable reducing the number of tests by automatically filling in the miss-

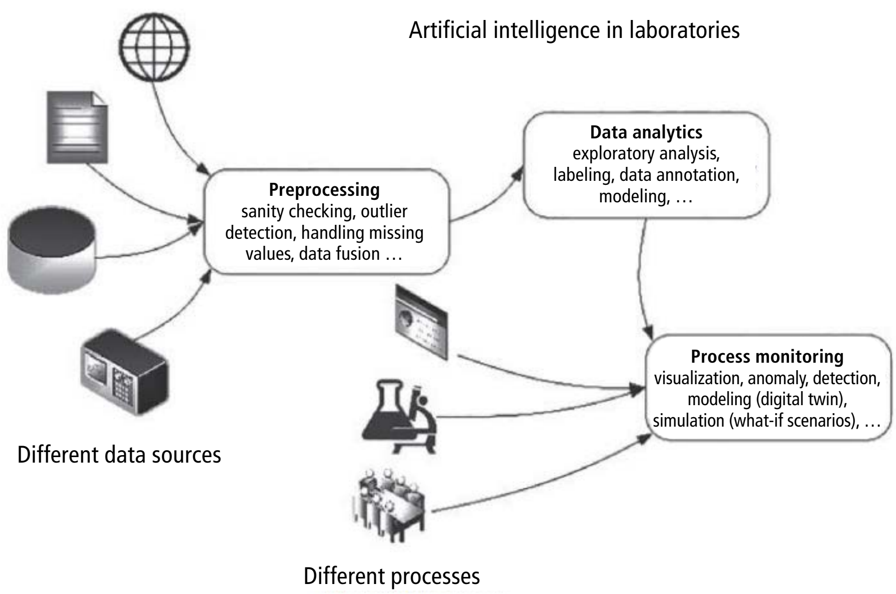

Figure 1. AI can support laboratories on different points, from data preprocessing and fusion of data from different sources to data analytics and monitoring of different processes. Source: Courtesy of Dunja Mladenic.

ing values or pointing out where more data should be collected.

AI methods today enable efficient organization of large amounts of heterogeneous data supporting efficient search and retrieval. On top of that, depending on the data modality, the user can be provided with powerful data exploration tools including rich data visualization, automatic outlier detection, data modeling, and prediction.

Moreover, data analytics can be applied in different phases of laboratory work from monitoring and directing data collection, data preprocessing, storing, and modeling to searching historical laboratory data (e.g., test results and notebooks) and enabling scientists to share data and models across problems and laboratories.

Process Monitoring

The processes in laboratories involve activities and data that may benefit from being monitored and modeled. We are talking about modeling dynamics and outcomes of a process or a set of possibly interconnected processes. Historical data can be used to build a reference model that can be adjusted, given the current context and trends of the monitored process.

As AI methods are utilized to construct a digital twin of a machine, logistic processes, or a production plant, they can also be used to construct a digital twin of some processes in the laboratories. This would enable monitoring for interdependencies, possible anomaly detection, and simulation of future developments of the process, giving the professionals the possibility to ask what-if questions.

With the capability of real-time modeling and monitoring of the incoming data, we can analyze digital laboratory notebooks over time within the same laboratory or across different laboratories. In a similar way, as machine learning methods have been used in decades of research publications to predict the next big thing in science,[1] one can analyze and monitor processes and data flows in laboratories.

Human in the Loop

Regardless of how much and in which processes in the laboratories we use AI, we should remember that AI could cover some of the intelligence, but there are other dimensions that humans bring to the process that cannot be covered by AI. We can say that intelligence is our capability to analyze some situations, learn from mistakes we have made or have seen others make, predict consequences of some action before taking it, and develop a strategy to achieve a goal. Intelligence is connected to clear, focused, and selective thinking and needs our guidance to avoid getting lost in details or fantasies. According to yogic philosophy, intelligence is one of the three creative forces needed to manifest success in life: consciousness, intelligence, and energy.[2] For instance, to know what and how we want to manifest, we need resources/energy to really do it; to know what to manifest and having resources, we need a strategy. Humans play a crucial role in the process by consciously deciding what is the goal of some work or laboratory experiments, intelligently developing a strategy for achieving the goal, and involving resources to implement the strategy and reach the goal.

References

[1] Lawton, G. The next big thing in chipmaking. Computer (Long. Beach. Calif). 40, 18–20 (2007).
[2] Mladenic, D. Artificial Intelligence (AI) Transforming Laboratories. in Digital Transformation of the Laboratory 289–295 (Wiley, 2021). doi:10.1002/9783527825042.ch21.

Active and Iterative Data-Dependent Reinjection Logic for Maintaining Throughput, Uptime, and Consistency in Triple Quadrupole LC/MS Analysis

The iReflex intelligent workflow tool maximizes analytical throughput and ensures samples are within tolerance.

Triple quadrupole LC/MS measurements are important for processing QA/QC samples involving:

- Pharmaceutical impurities
- Pesticides in food and the environment
- Veterinary drug detection in foods
- Measurement of biological analytes

Key objectives are:

- Consistent results
- Avoidance of sample reprocessing
- High sample throughput

What is iReflex?

The 6475 triple quadrupole LC/MS system with MassHunter 12 features an intelligent workflow called iReflex.

iReflex uses an active and immediate data processing algorithm that evaluates and reinjects samples in a data-dependent manner based on:

- Detection of a previous sample carryover
- Detection of a sample outside of the calibration range
- Fast analyte screening

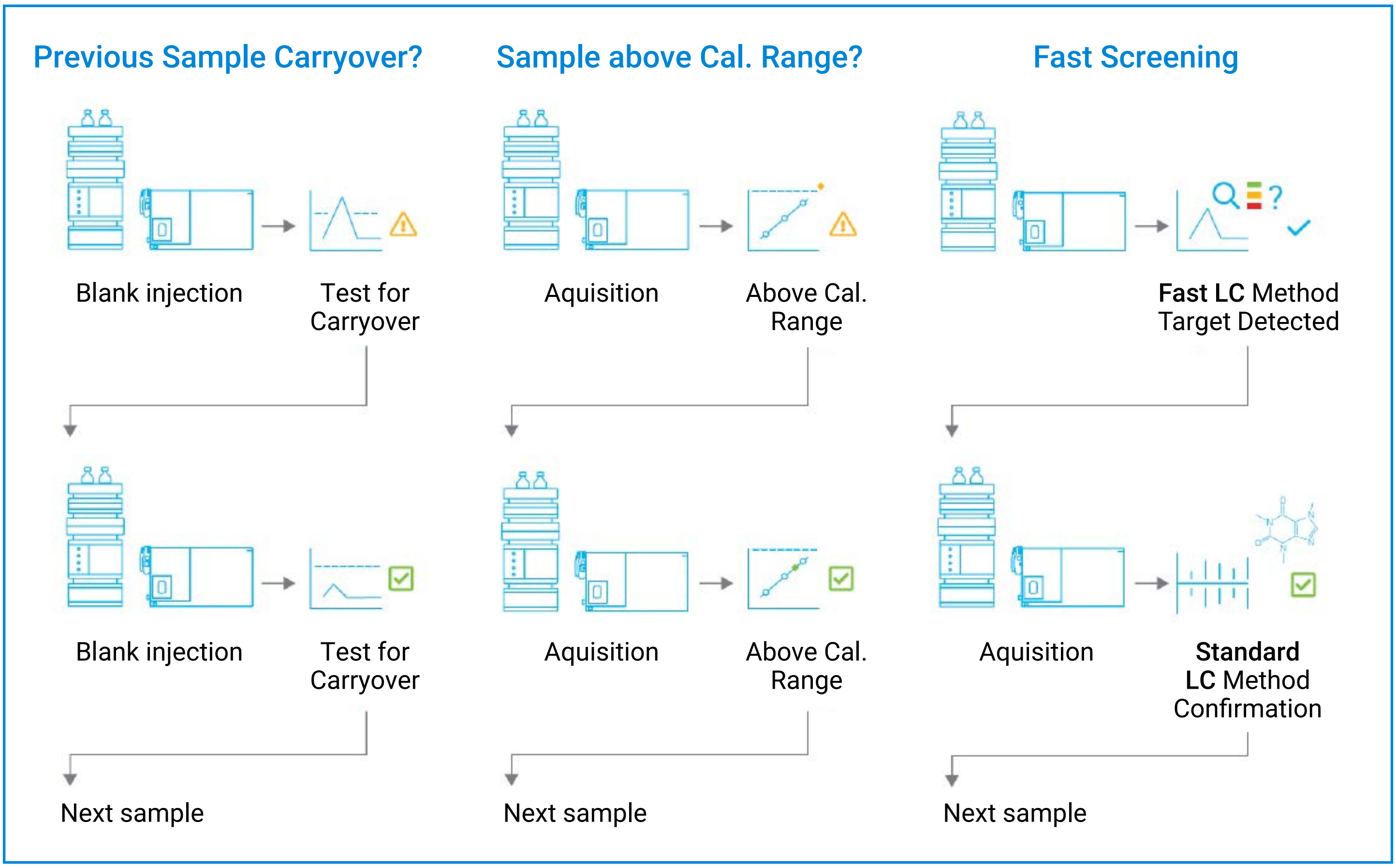

Figure 1: Inteligent Relex (iReflex) Workflow logic.

Key features of iReflex

iReflex workflows:

- Boost lab productivity by reducing manual intervention and sample reprocessing
- Save valuable sample material by automatically preventing contamination
- Enhance throughput for large batch analysis through automation
- Automatically generate a combined report

Detection of carryover in a blank above the outlier threshold triggers the insertion of blanks, with the additional option to pause the worklist to prevent contamination of samples.

If a target analyte concentration is within the calibration curve, iReflex triggers an insert/append re-injection with reduced volume to provide an estimated concentration.

How does iReflex speed up analysis?

Fast screening methods help increase sample throughput by rapidly identifying presumptive positive samples, which are then manually scheduled for reinjection and analysis.

iReflex automates reinjection and analysis of presumptive positive samples using a different analysis method.

The fast screening iReflex workflow produces two different data batches:

Original worklist with the fast-screening method

1st tier

Reinjected samples analyzed using a more comprehensive method

2nd tier

Three iReflex workflows can operate concurrently within one worklist to ensure samples are measured within SOP guidelines.

Accelerated Lifetime Testing with Real-Time Early Maintenance Feedback (EMF) Diagnostic Monitoring on the 6475 Triple Quadrupole LC/MS

How can you verify instrument health and longevity in a triple quadrupole LC/MS system?

With use, ion optics components in a triple quadrupole LC/MS system can become soiled, leading to concerns about stability.

An experiment tested the new Agilent 6475 triple quadrupole LC/MS system, which has onboard intelligence to report on instrument health and status through Early Maintenance Feedback (EMF).

The experiment

Bovine urine was diluted 1:1 in acetonitrile/water and delivered to the system using an Infinity II 1290 HPLC with a dual injector setup in overlapped injection mode with isocratic flow of 90:10 acetonitrile/water + 0.1% Formic acid.

A ZORBAX Extend-C8, 80Å, 2.1 mm, 1.8 µm, 1,200 bar pressure limit, UHPLC guard column was used to:

- Generate sufficient backpressure for stable HPLC operation
- Simulate the use of an analytical HPLC column

MRM signals of various analytes were recorded to 'age' the Electron Multiplier horn as if it were in use.

EMF provided real-time monitoring of instrument performance including:

- Ion injector blockage
- Precipitation on the nebulizer
- Detector estimated lifetime
- Disruptive events

An injection series was carried out; then the ion source and desolvation assembly were examined to check for:

- Ion burn
- Salt accumulation
- Broad matrix deposition
- Potential modes of failure

Check tunes were performed every 1,000 injections with no cleaning of the nebulizer or ion injector.

Results

Despite heavy front-end contamination, mass calibration (*m/z* drift) and mass spectral peak width (FWHM) remained within tolerance and stable over 10,000 injections.

No critical EMF events were triggered throughout the investigation.

Despite the constant bombardment of ions and a notoriously heavy matrix:

- Detector health was stable
- Nebulizer and ion injector remained unclogged
- Spray stability remained consistent

Development and Validation of a Quantitative Method for Multiresidue Pesticides in Food Matrices Using the Agilent 6475 LC/TQ System

How well does the new Agilent 6475 LC/TQ system perform in pesticide analysis?

Sensitive and robust quantification of pesticide residue in food matrices is required to manage food safety and environmental impact.

The new triple quadrupole LC/MS system 6475 LC/TQ has several improved features:

- Scheduled autotune and early maintenance feedback
- Enhanced software for method optimization
- iReflex reflexive reinjection logic
- Supporting Title 21 CFR Part 11 and Annex 11 compliance

The experiment

A 1290 Infinity II Bio LC system coupled with a 6475 LC/TQ was used to quantify 497 pesticides in food matrices including organic wheat, olive oil, and black tea.

The pesticide mix was spiked into solvent and food extracts to make matrix-matched standard curves with concentrations ranging from 0.1 to 50 µg/L.

Samples were analyzed with a dynamic MRM-based LC-MS/MS method using MassHunter Workstation for LC/TQ 12.0.

The method used monitored 1003 MRMs within a 20-minute LC gradient, which shows good chromatogram separation and improved lab productivity.

Results

The method performance showed excellence in:

- Linearity with 92% pesticides in wheat, 94% in olive oil, and 92% in black tea showing R2 > 0.99
- Precision and accuracy at all calibration levels
- Sensitivity of 94% of pesticides in wheat, 98% in olive oil, and 94% in black tea show a lower limit of quantification LLOQ) ≤ 10 µg/kg

Outstanding reproducibility was shown for RT and MS signal in 5 days of continuous running with a black tea extract spiked with pesticide standards at 10 µg/kg