AI SUMMIT

CINCINNATI, OH · NOVEMBER 14–16, 2023

Building Quality Science using AI

Prof. Dr. Thomas Friedli

Matteo Bernasconi

AFDO RAPS HEALTHCARE PRODUCTS COLLABORATIVE

Inspiring Collaboration. Leading Innovation. Making a difference.

Funding for some of the research contained in this presentation is made possible, in part, by the Food and Drug Administration through BAA contracts [#75F40121C00161, #75F40119C10121] and grants [#5U01FD005675-02, #1U01FD005675-01]. The views expressed in written materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does any mention of trade names, commercial practices, or organization imply endorsement by the United States Government

Agenda

1. FDA BAA: RiskSurve – Overview & Results
2. Our Data and Scientific Background
3. Summary & Outlook

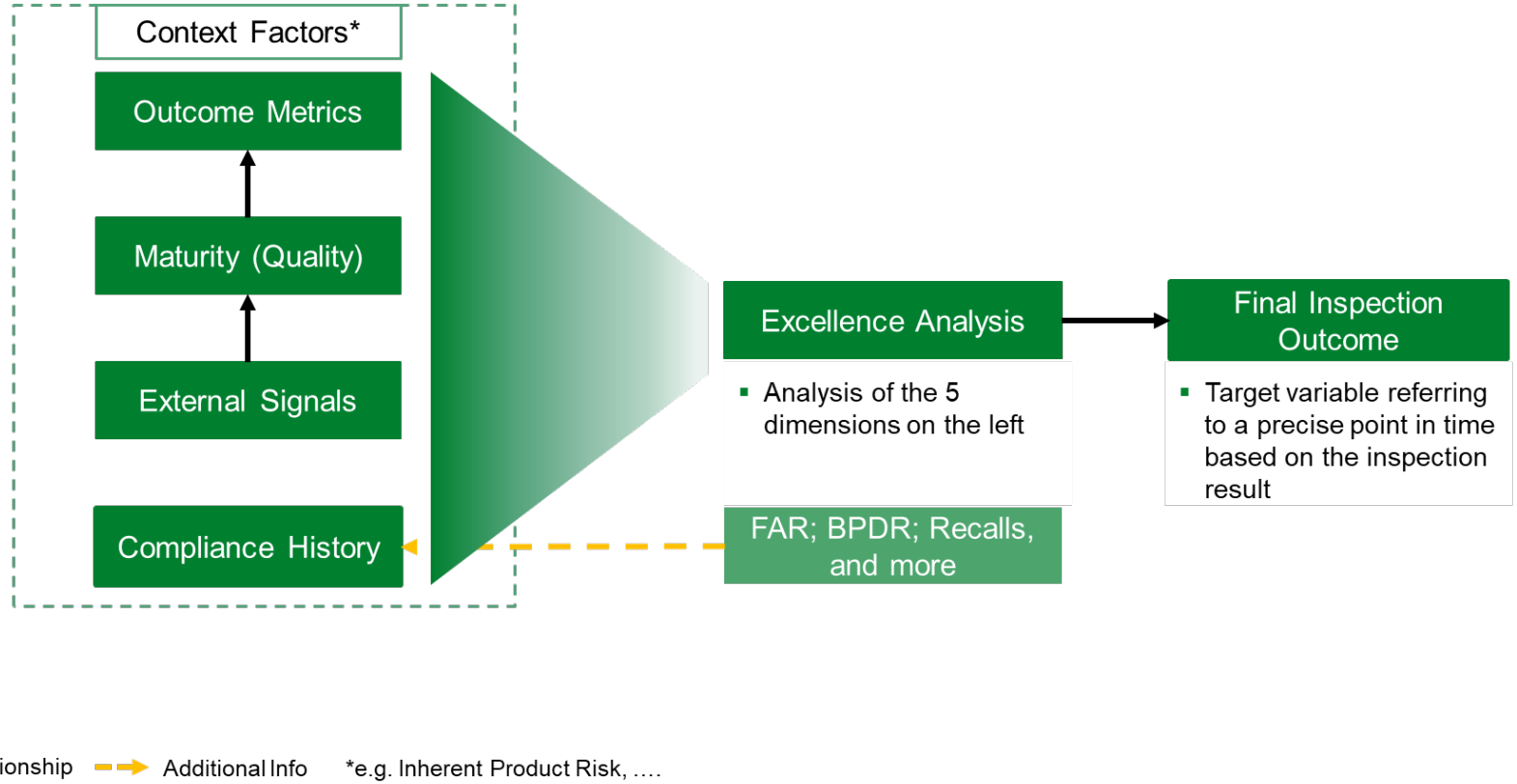
Agenda

1. **FDA BAA: RiskSurve – Overview & Results**
2. **Our Data and Scientific Background**
3. **Summary & Outlook**

FDA's ongoing efforts to characterize site quality in the context of a broader surveillance strategy motivated the goal of our current BAA

This project aims to create a comprehensive Remote Site Risk Surveillance Model consolidating data from the four dimensions Outcome Metrics, Quality Management Maturity, Compliance History, External Signals embedded in their relevant Context.

RiskSurve relies on a conceptual framework to drive our analysis and develop the predictive model



We operationalized four dimensions

Outcome Metrics – 9 Metrics in 4 categories

- ***Maintenance***: *e.g. Unplanned Maintenance*
- ***Quality***: *e.g. Rejected Batches*
- ***Delivery***: *e.g. On Time In Full*
- ***Efficiency***: *e.g. Maintenance FTEs/ Overall FTEs*

Compliance History – 2 Perspectives

Site Perspective:

- *Past compliance Information from the manufacturing facility*

Corporate Perspective:

- *Identification of network wide quality failures to rise flag alerts related to the manufacturing network*

Maturity (Quality) – 13 Items in 3 categories

- ***Performance Measurement & Continuous Improvement***
- ***Collaboration Culture & Organization***
- ***Training & Skills***

External Signals – Proxies for Performance & Maturity

Proxy for Performance & Maturity on Corporate Level:

- *Collection of publicly available data from the web to calculate proxies for:*

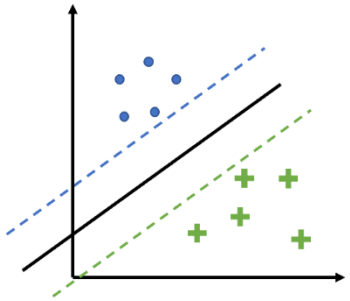
Employee Culture – *Site vs. Corporate, Low vs. High tiers*

Performance & Complexity – *Financial & Product related information*

We have tested three Classification Models

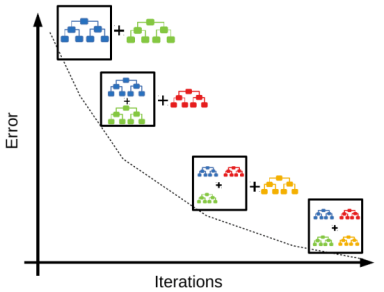
1 Support Vector Machines

- Find an optimum decision boundary that separates datapoints belonging to different classes
- Efficient for small datasets, handles multi-modality, does not get stuck in local minimum



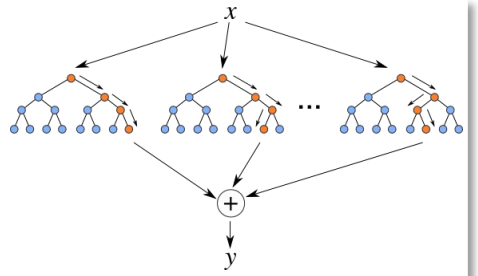
2 LightGBM

- A boosting framework that uses ensemble of decision trees similar to XGBoost; gradient boosting of performance on model residuals
- Faster training speed, better accuracy, lower memory usage, handles bigger datasets



3 Random Forest

- Based on bagging of decision tress on randomized bootstrapping of data and random subset of features
- Faster processing speed, useful for ranking variable importance, but more importance on hyperparameter for model performance optimization



After comparing the accuracy of three different classification models, we selected the Light Gradient Boosting Machine (LightGBM)

During the model development, we observed **how tree-based** models (LightGBM and Random Forest) **outperformed** the Support Vector Machine model. The **quicker computation time and accuracy** of the result made us **selecting the LightGBM** over the RF

LightGBM Multiclass Classification

	Dimension & Context Factors			
	$Accuracy_{avg}$	$F1_{NAI}$	$F1_{VAI}$	$F1_{OAI}$
FIR0	56%	40%	70%	0%
FIR1	61%	50%	71%	0%
FIR2	**65%**	**50%**	**74%**	**57%**

LightGBM Sequentially Binary Classification

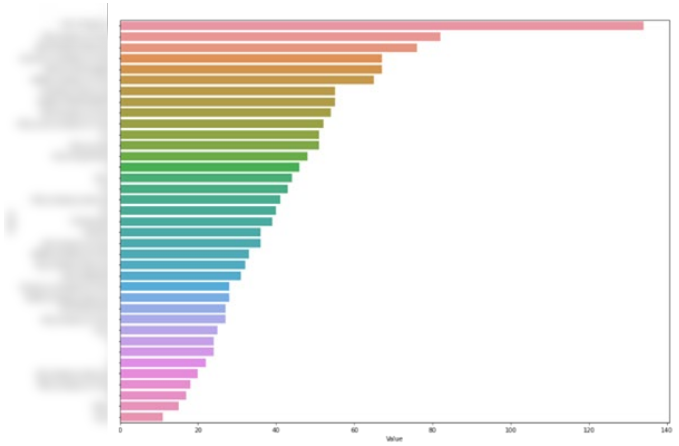
	Step 1: NAI vs VAI & OAI			Step 2: VAI vs. OAI		
	$Accuracy_{avg}$	$F1_{NAI}$	$F1_{Rest}$	$Accuracy_{avg}$	$F1_{VAI}$	$F1_{OAI}$
FIR0	56%	33%	67%	75%	84%	40%
FIR1	61%	40%	71%	70%	82%	0%
FIR2	**57%**	**27%**	**69%**	**88%**	**92%**	**67%**

Results in multiclass and binary classification settings showed better accuracy with FIR2, since the model has more data available. Additionally, due to our aim, the sequentially binary classification is more suitable and reveals better accuracy compared to the multiclass.

Therefore, we continued with a sequentially binary classification with FIR2 as target variable.

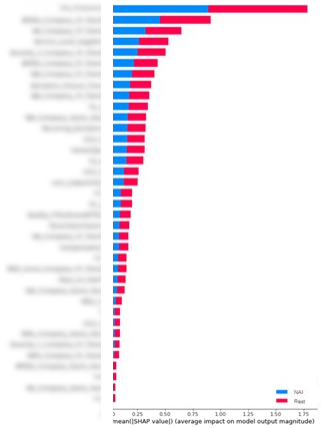
Overview on the feature selection in iteration 4 Features Overview Overall Model

Feature Selection



	Number of Features 37(70)*
Outcome Metrics	5 (9)
Maturity (Quality)	7 (16)
Compliance History	16 (28)
External Signals	9 (17)

Feature Importance

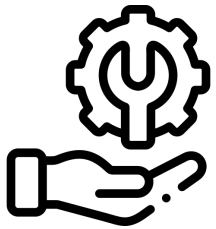


- Compliance history is the dimension with the majority of features in this iterations. Site perspective is the feature with the greater contribution;
- However, for better model accuracy, a balanced mix between the four dimensions is required

*Number not in brackets showing the number of features in this model/iteration. Number in brackets showing the overall number of features in the initial model.

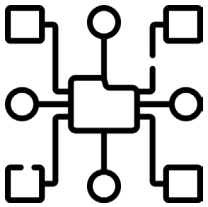
Building upon existing results we needed additional qualitative insights

We aimed at proposing additions to the site surveillance strategy



Technically

Inlucding addtional influential factors with critcality levels to existing site selection algorithm.



Organizationally

How to include the additional influential factors and what are the implications to the site selection process.

… by leveraging findings from year-one and new insights

1 Findings from year-one

The basis for year two extension. Identified important relations between compliance history, maturity, outcome performance, and external signals.

2 Criticality Levels for selected Metrics

Derivation of criticality levels, upper or lower limits, to flag risks. Provides information about the influential degree and boundary conditions of parameters.

3 Site Excellence Ranking Logic

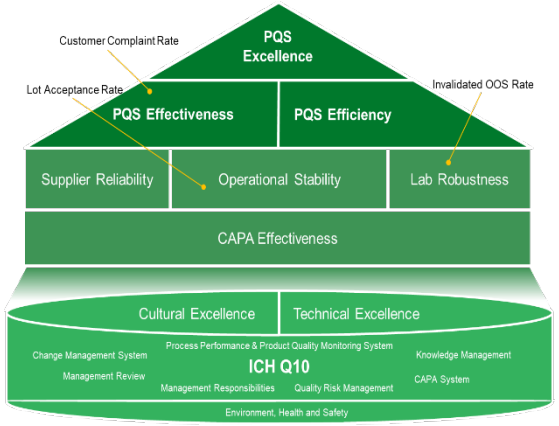
Derivation of a site ranking logic and measurement scale. Categorization might have an impact on surveillance strategy.

4 Qualitative Validation

Interviews with regulators (FDA & PICS) as well as with **the industry**

Our Excellence Score should consider several perspectives

PPSM



Sand Cone Model

Steps to QMM

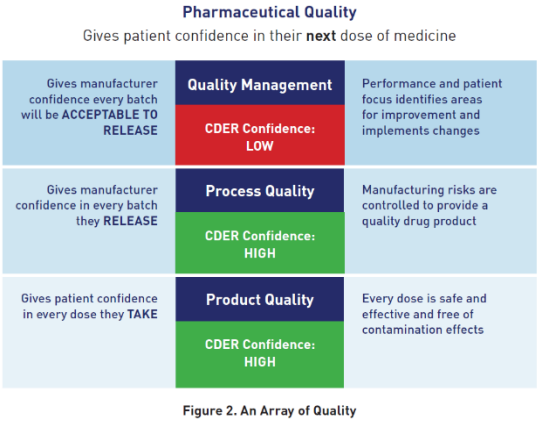
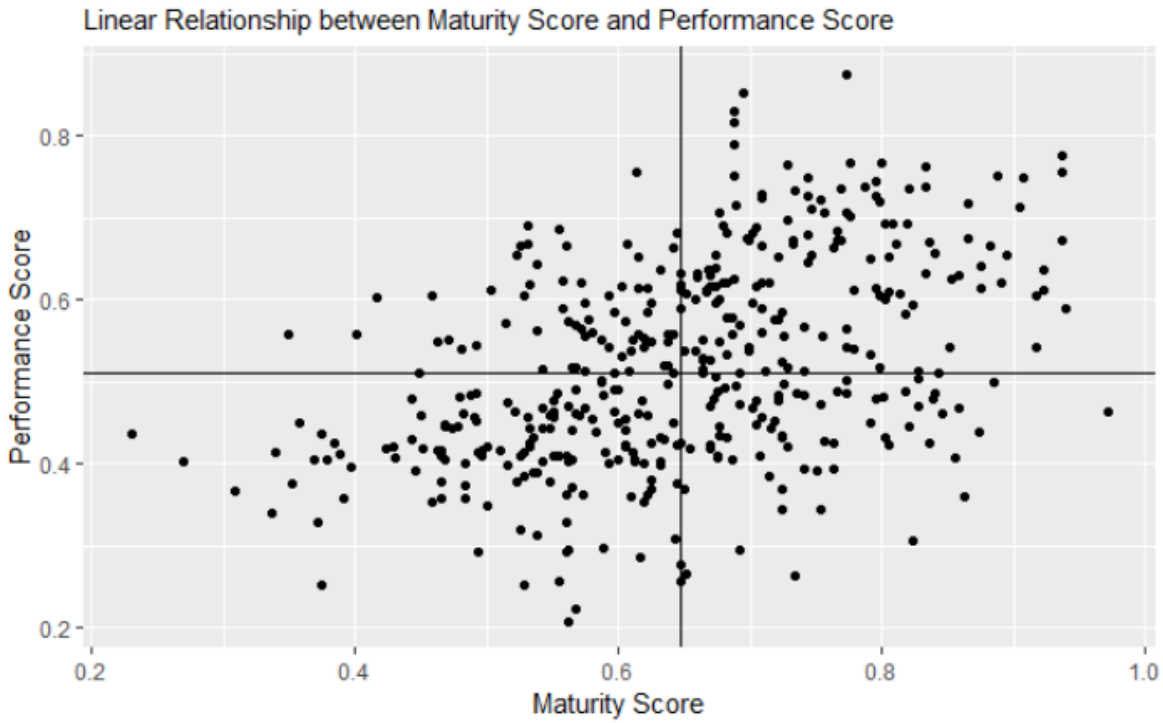


Figure 2. An Array of Quality

FDA (2022). Quality management maturity: essential for stable U.S. supply chain of quality pharmaceuticals
Ferdows & de Meyer (1990). Lasting improvements in manufacturing performance: in search of a new theory. Journal of Operations Management
Friedli et al. (2019). FDA quality metrics initiative – third year report.

The excellence score must consider both performance and maturity scores simultaneously, since the analysis reveals the existence of synergies



Our analysis shows a positive linear relationship between maturity and performance. The Excellence Score must reflect this relationship and especially the 4 quadrants depicted in the graph, by providing a higher weighting to maturity instead of performance since this provides the basis for a sustainable performance outcome.

Our concept of the Excellence Score must respect multiple criteria to guarantee a correct scoring and resulting ranking logic – we tested two possibilities

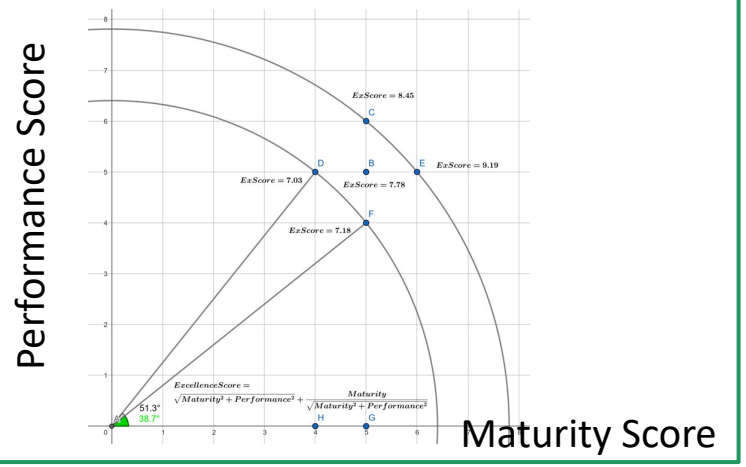
EXCELLENCE SCORE

Criteria

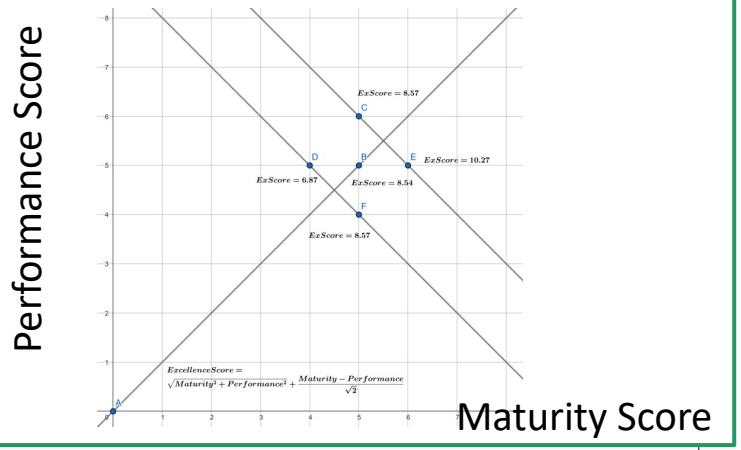
1) For equal performance score a higher maturity score must lead to a higher Excellence Score;
2) For equal maturity score a higher performance score must lead to a higher Excellence Score;
3) If the average value of performance and maturity score is the same for two establishment, the one with the higher maturity score must have a higher Excellence Score;

Options

A) Angle-based Correction



B) Orthogonal-based Correction



Agenda

1. FDA BAA: RiskSurve – Overview & Results
2. Our Data and Scientific Background
3. Summary & Outlook

Our competitive advantage? The availability of data!

Outcome Metrics – 9 Metrics in 4 categories

- ***Maintenance**: e.g. Unplanned Maintenance*
- ***Quality**: e.g. Rejected Batches*
- ***Delivery**: e.g. On Time In Full*
- ***Efficiency**: e.g. Maintenance FTEs/ Overall FTEs*

Maturity (Quality) – 13 Items in 3 categories

- ***Performance Measurement & Continuous Improvement***
- ***Collaboration Culture & Organization***
- ***Training & Skills***

Compliance History – 2 Perspectives

***Site Perspective**:*

- *Past compliance Information from the manufacturing facility*

***Corporate Perspective**:*

- *Identification of network wide quality failures to rise flag alerts related to the manufacturing network*

External Signals – Proxies for Performance & Maturity

***Proxy for Performance & Maturity on Corporate Level**:*

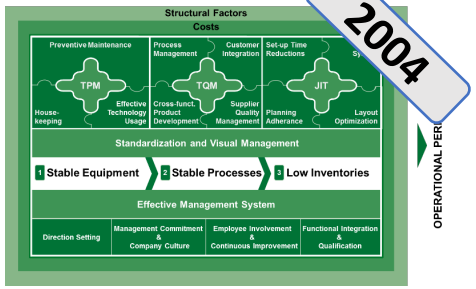
- *Collection of publicly available data from the web to calculate proxies for:*

***Employee Culture** – Site vs. Corporate, Low vs. High tiers*

***Performance & Complexity** – Financial & Product related information*

Our data have been collected over the years in three main benchmarking exercises

St.Gallen OPEX Model

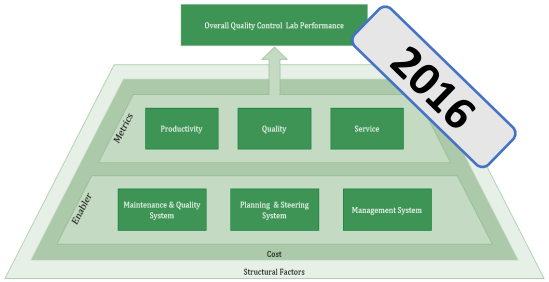


 # +400 Establishments

Manufacturing

Potentially 9'962 data points per site

QC Excellence Model

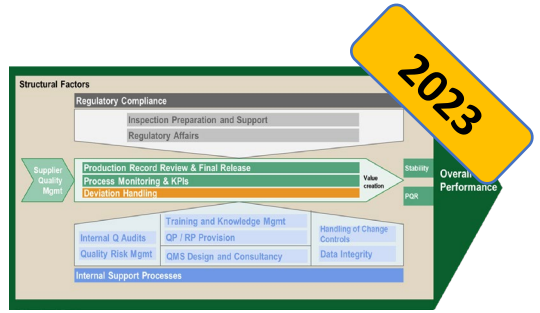


 # +130 QC Labs

Quality Control Laboratories

Potentially 356 data points per Lab

QA Excellence Model



 # 13 QA Functions

Quality Assurance Functions

Potentially 453 data points per Function

We have a solid backbone of operational data to empirically investigate multiple questions!

We conduct our analysis by following some well known theories

vs.

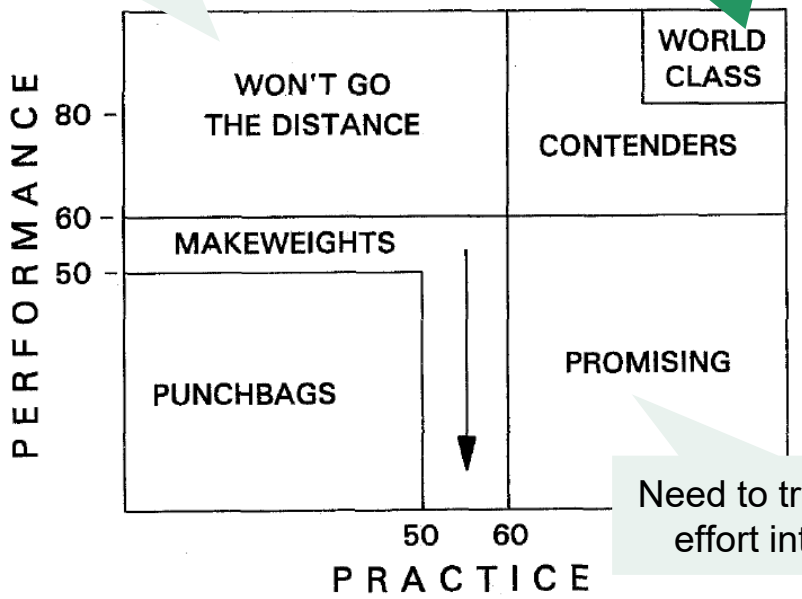
Source: Juan Andres at ISPE/FDA/PQRI Quality Manufacturing Conference in Washington DC (2015); Ferdows & de Meyer (1990)

Sustainable performance improvements can only be built on maturity – in systems, processes, tools, and people

Theoretical foundation

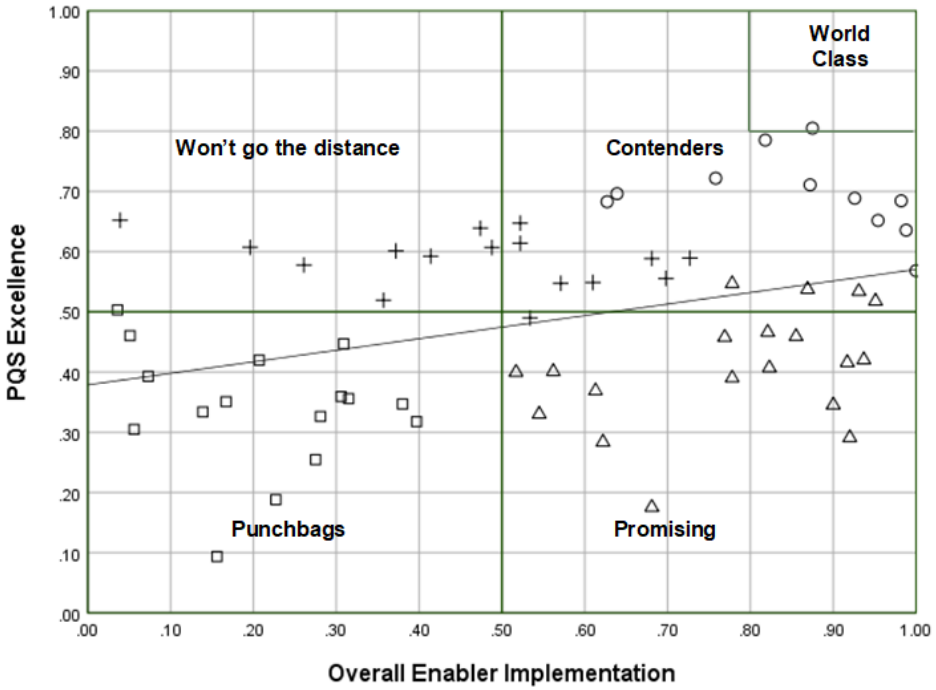
not sustainable, short-term wins

Desired state



Need to transfer past effort into results

Practical Reality from our Projects



Source: Voss et al. (1995)

Agenda

1. FDA BAA: RiskSurve – Overview & Results
2. Our Data and Scientific Background
3. Summary & Outlook

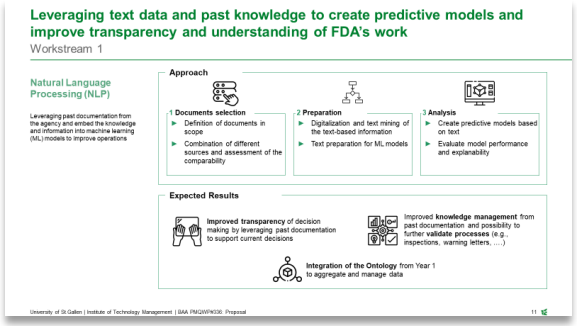
Summary

1. A combination of regulatory and operations data yields the best results in predicting final inspection classifications
2. An ontology supported us in data management, data selection, and made the later model outputs more open to interpretation
3. Data is key
4. The potential of AI to deepen our understanding about interdependences is enormous
5. AI has the potential to support the development of scientific base for quality in operations also beyond pharma

We have been awarded a new 2-years BAA with the FDA

Drive Predictive Continuous Improvement (CI) Acceleration – Towards Performance Based Regulation Regulations and Oversight

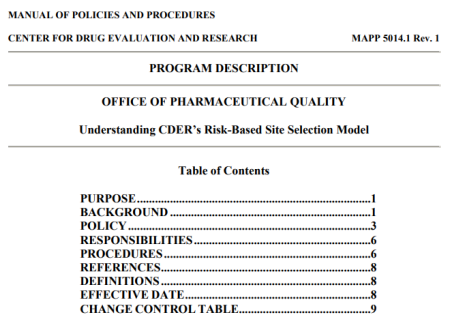
FDA Knowledge Management System



Work Stream Expected Results

- Improved transparency of decision-making process
- Improved knowledge management
- Further process validation
- Integration into the ontology from Y1 RiskSurve Project

Site Selection Model Review


MANUAL OF POLICIES AND PROCEDURES

CENTER FOR DRUG EVALUATION AND RESEARCH MAPP 5014.1 Rev. 1

PROGRAM DESCRIPTION

OFFICE OF PHARMACEUTICAL QUALITY

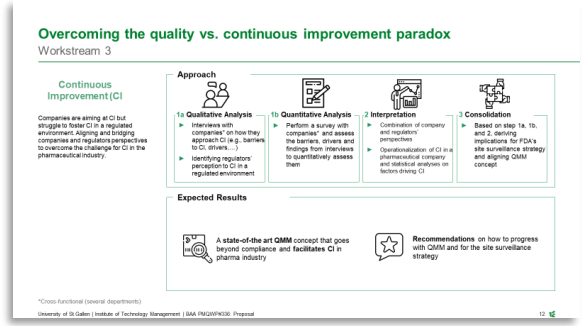
Understanding CDER's Risk-Based Site Selection Model

Table of Contents

Work Stream Expected Results

- Strengths and Weaknesses of current approach
- Evaluation of possible additional metrics and aggregation logics
- Update recommendations

Continuous Improvement



Work Stream Expected Results

- QMM facilitates continuous improvement in pharma industry
- Recommendation on how to overcome CI hurdles
- Refinement of predictive models
- Considerations of the integration of QMM into site surveillance strategy

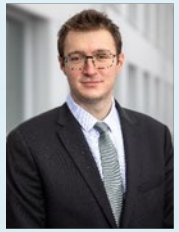
Contact Details

Please do not hesitate to contact us if you have any questions

Prof. Dr. Thomas Friedli

Director
Institute of Technology Management (ITEM-HSG)

Dufourstrasse 40a
CH - 9000 St.Gallen



Matteo Bernasconi

Group Coordinator

matteo.bernasconi@unisg.ch
+41 (0)71 224 72 71