



HAL
open science

Decision support for the comparative evaluation and selection of analytical methods: detection of genetically modified organisms as an example

David Dobnik, Kristina Gruden, Jana Zel, Yves Bertheau, Arne Holst-Jensen,
Marko Bohanec

► To cite this version:

David Dobnik, Kristina Gruden, Jana Zel, Yves Bertheau, Arne Holst-Jensen, et al.. Decision support for the comparative evaluation and selection of analytical methods: detection of genetically modified organisms as an example. Food Analytical Methods, 2018, pp.1-18. 10.1007/s12161-018-1194-1 . hal-02623338

HAL Id: hal-02623338

<https://hal.inrae.fr/hal-02623338>

Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Decision Support for the Comparative Evaluation and Selection of Analytical Methods: Detection of Genetically Modified Organisms as an Example

David Dobnik¹ · Kristina Gruden¹ · Jana Žel¹ · Yves Bertheau^{2,3} · Arne Holst-Jensen⁴ · Marko Bohanec⁵

Received: 8 August 2017 / Accepted: 5 February 2018
© The Author(s) 2018. This article is an open access publication

Abstract

The selection of the best-fit-for-purpose analytical method to be implemented in the laboratory is difficult due to availability of multiple methods, targets, aims of detection, and different kinds and sources of more or less reliable information. Several factors, such as method performance, practicability, cost of setup, and running costs need to be considered together with personnel training when selecting the most appropriate method. The aim of our work was to prepare a flexible multicriteria decision analysis model suitable for evaluation and comparison of analytical methods used for the purpose of detecting and/or quantifying genetically modified organisms, and to use this model to evaluate a variety of changing analytical methods. Our study included selection of PCR-, isothermal-, protein-, microarray-, and next-generation sequencing-based methods in simplex and/or multiplex formats. We show that the overall result of their fitness for purpose is relatively similar; however, individual criteria or a group of related criteria exposed more substantial differences between the methods. The proposed model of this decision support system enables easy modifications and is thus suitable for any other application of complex analytical methods.

Keywords Multicriteria decision analysis · Genetically modified organisms · Method evaluation · DEXi · Decision support system · DSS

Introduction

Analytical laboratories have to select and implement new methods from time to time. Alternatively, based on received real-life samples, they may have to select among a panoply of methods for multiple target analyses. This selection can be rather arbitrary, primarily directed by tradition or the expertise

of trained persons, but it can also be based on systematic objective and reproducible comparative analysis. Options of choice often include the use of different technological platforms and the choice between simplex and multiplex methods. With increasing numbers of optional methods, the complexity of selecting analyses will either require a long reflection period of trained analysts or some sort of decision support system, selection criteria, and rules for comparison of attributes as a way to reduce uncertainties and ensure objective and reproducible choices independently of the presence of trained experts. In many fields where analytical methods are applied, the situation is quite complex. In this manuscript, we take detection (including identification and quantification) of genetically modified (GM) organisms (GMOs) as an example to illustrate how the decision process can be facilitated in a complex situation through the use of decision support system (DSS).

Many countries around the world regulate cultivation and trade with GMOs and have implemented an authorization system and mandatory labeling above a certain threshold (Grùère and Rao 2007; Vigani et al. 2012; Milavec et al. 2014). For example, in the European Union (EU), the labeling threshold

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12161-018-1194-1>) contains supplementary material, which is available to authorized users.

✉ David Dobnik
david.dobnik@nib.si

- ¹ Department of Biotechnology and Systems Biology, National Institute of Biology, Večna pot 111, Ljubljana, Slovenia
- ² Institut national de la recherche agronomique, Versailles, France
- ³ Present address: Muséum national d'histoire naturelle, Paris, France
- ⁴ Norwegian Veterinary Institute, Oslo, Norway
- ⁵ Jožef Stefan Institute, Ljubljana, Slovenia

for food products containing, consisting of, or produced from authorized GMOs is set at 0.9% by Regulation (EC) 1829/2003 (European Commission 2003). After more than two decades of commercialization of GM plants, the number of new GM events and the diversity of GM traits and constructs, as well as the mixture of GMOs and of their derived products and the purpose of analyses (labeling vs. market withdrawal), is still steadily increasing (James 2015). The complexity of the products that need to be tested for the presence of GMOs is therefore also increasing. This complexity will increase even more, if products derived from new breeding techniques would also have to be labeled. International trade and differences in national regulations combined with the need for harmonized testing and results interpretation add to the complexity of the method selection in the analytical laboratories.

Currently, detection, identification, and quantification of GMOs (Holst-Jensen et al. 2012) are mostly done using quantitative real-time polymerase chain reaction (qPCR). The individual methods either target a DNA motif present inside the inserted GM construct or a junction between the hosts' native DNA and the inserted sequence (see, e.g., Holst-Jensen et al. (2012) for a detailed review). Screening for the presence of genetic elements commonly found in GMOs can be an effective approach to reduce the number of needed analyses (Holst-Jensen et al. 2012). It is, however, prone to misinterpretation and normally will have to be complemented with the use of methods targeting specific GMOs (transformation event-specific methods) (Holst-Jensen et al. 2012). Multiplex screening approaches can improve the cost efficiency (Huber et al. 2013) but also have important limitations. Several more or less specific PCR (Randhawa et al. 2009, 2010; Holck et al. 2010; Luan et al. 2012) and qPCR (Germini et al. 2004; Waiblinger et al. 2007; Bahrtdt et al. 2010) multiplex methods have already been developed. Additionally, alternative testing methods and targets are available (reviewed in Milavec et al. (2014)), and new methods and technologies/platforms are constantly developed.

Altogether, the GMO detection and identification processes should be reproducible, understandable, transparent, and user-friendly for both the analyst and the decision maker. Selecting among analytical strategies to meet these objectives is challenged by the complexity of the methods. The integrated use of additional sources of information would be beneficial, as would the possibility to weigh issues like personnel training costs and restricted budgets.

Factors considered for the selection of each method are as follows: properties of the product (product type, ingredients, diversity of targets that may have to be detected, etc.), purpose of the analysis (screening, identification, quantification, as well as unapproved product, detection for market withdrawal, etc.), performance parameters of the methods (limit of detection, applicability to the situation, compatibility with other methods in the sequence, etc.), capabilities of the laboratory

(available equipment, skilled personnel and other resources, reference materials, etc.), duration of analysis in case of urgency, practicability, cost of setup, and running costs. Consequently, this proves to be a complex decision problem, not restricted to analytical methods, which requires extensive knowledge and experience.

The assessment and selection of analytical methods involves a multitude of criteria, which are generally conflicting and affect the decision in different ways. Until now, there are some reports on comparisons of DNA extraction methods, where performance of individual parameters was compared and statistically evaluated (Jasbeer et al. 2009; Volk et al. 2014). However, no overall assessment of all parameters together was performed. The problem of including many parameters and criteria in the evaluation can be defined as a multicriteria decision problem. It is unlikely that one method will be the best for all of the considered criteria. Thus, approaches implementing multicriteria decision analysis (MCDA) (Ishizaka and Nemery 2013; Greco et al. 2016) may prove useful. In MCDA, each alternative (method) is first assessed according to each criterion. These individual assessments are then aggregated into an overall evaluation of each alternative method. On this basis, the alternatives are compared and/or ranked, and the best one can be chosen. Various analyses, such as sensitivity analysis and "what-if" analysis, are also possible to further support and justify the assessment. Pakpour (2012) reported the use of MCDA with weighted sum method for evaluation of sample preservation coupled with DNA extraction methods. The reported system, however, included a rather small amount of attributes in the study. Another tool named Analytical Method Performance Evaluation (AMPE), implementing more computational analysis, was developed to help with analytical method validation, including data handling capabilities and series of statistical calculations (Acutis et al. 2007). This tool is thus suitable only for comparisons of different parameters (e.g., limits of quantification and detection, variability, accuracy, trueness, precision) of one analytical method, produced by different users, which is usually the case in collaborative method validations. Currently, there are no tools available that would help in a thorough, objective (human independent), and reproducible evaluation and comparison of several analytical methods from different platforms (e.g., PCR, next-generation sequencing [NGS], and protein-based methods), taking into account a multitude of criteria that affect the methods' performance and applicability.

This study addresses a common situation faced by enforcement or private analytical laboratories and other organizations involved in the analytical traceability and controls in the food/feed production chains (Holst-Jensen 2009). Numerous traceability data (country of origin, species, industrial process, etc.) are available and processed, before a sample is submitted to the analytical laboratory. These data can obviously be used to

direct the analyses, i.e., choosing the appropriate sampling strategy and the required detection process (routine or more targeted analysis). All these data can be integrated in MCDA by the analyst for detection, results interpretation, or decision-making (Bohanec et al. 2013a).

Although MCDA could also be applied to additional steps, such as sample processing and DNA extraction, we have here limited our focus to the analysis of purified DNA from the sample(s). DNA analysis is usually carried out as a sequence of methods, most commonly involving a combination of screening, identification, and quantification methods.

The aim of our work, as part of a more general work on the development of GMO detection and identification strategies, was twofold: (1) to prepare an MCDA model suitable for evaluation and comparison of analytical methods and (2) use the model to evaluate a variety of analytical methods taking the field of GMO detection and/or quantification as an example case. Methodologically, we built upon the results of the European FP6 project Co-Extra (Bertheau 2013). There, as part of the Co-Extra decision support system (Bohanec et al. 2013a), a number of MCDA models were developed, including two models for the assessment of analytical methods. For the purpose of this study, we have, within European FP7 Decathlon project (www.decathlon-project.eu), adapted one of the models, called AM_DetQuant (Bohanec et al. 2013a). We have narrowed the original model to only assess whether a method is fit for purpose (*FitForPurpose*) in the situation being evaluated and to determine, which method is best for purpose (*BestForPurpose*). Rather than focusing on the comparison of several different methods for qPCR, our aim was to prepare and use the multicriteria model for comparing and evaluating a variety of different platforms used for running analytical methods. Altogether 15 different methods were lined up and evaluated from different aspects.

For clearer presentation, we prepared decision rules in a way that would allow all of these methods to be compared side by side, independently on their position in the whole pipeline of GMO analysis. In current GMO testing, several methods are usually combined and performed sequentially. The idea behind the selected decision rules was that they enable comparison and selection of methods for the same purpose (e.g., quantification), rather than implementing rules for more complex scenarios where methods for different purposes, such as screening and event-specific quantification, are combined. This also means that the exercise must be repeated for each desired purpose, with purpose-based decision rules.

In this manuscript, we show that the MCDA model makes possible the direct comparisons of several unrelated technologies. Notably, even when their overall fitness for purpose is relatively similar, comparing and evaluating individual criteria or a group of related criteria can uncover substantial differences between methods and technologies. The newly

developed model enables easy modifications of the criteria and of their influence on final evaluation. Thus, it can be easily adapted to any other complex analytical situation for selecting the most suitable analytical method(s). Particularly, it could be of great help in highly complex situations, where results of different identification techniques and approximate data, such as NGS sequences, would have to be combined with traceability elements.

Materials and Methods

Analytical Methods Assessed in This Study

Methods, selected for evaluation and comparison (assessment) in this study, are listed in Table 1. Four of the methods represent the qPCR system currently used in routine GMO diagnostics covering different applications: simplex and multiplex screening/identification (Alary et al. 2002; Kuribara et al. 2002; Pla et al. 2013) and simplex quantification (Holck et al. 2002). Two additional qPCR applications (SIMQUANT; Berdal et al. 2008) use qPCR chemistry together with the limiting dilutions principle, which is near to the idea of the ddPCR-based methods, of which two were included (Morisset et al. 2013; Dobnik et al. 2015). Other selected methods include LAMP with end-point fluorescent (Chen et al. 2011; Wang et al. 2015) or bioluminescence real-time detection (Kiddle et al. 2012), multiplex PCR with hybridization on microarrays (Leimanis et al. 2008; Hamels et al. 2009) or detection with capillary gel electrophoresis (Nadal et al. 2006), a protein-based method (Van Den Bulcke et al. 2007), and two NGS methods (unpublished, developed within the EU FP7 Decathlon project), one for enriched samples and another for whole genome sequencing (see, e.g., Arulandhu et al. 2016; Holst-Jensen et al. 2016). The majority of the selected methods are validated in-house or within international collaborative trials and their fitness for purpose demonstrated elsewhere (see Table 1 for references).

Qualitative MCDA Method DEX

The original AM_DetQuant (Bohanec et al. 2013a) model, which was adapted for the purpose of this research, was developed using an MCDA method DEX (Decision EXpert) (Bohanec et al. 2013b; Bohanec 2015). DEX is a qualitative MCDA method, specifically designed to support expert modeling, that is, acquisition of decision knowledge from experts and decision makers in the form of decision rules. As all other MCDA methods, DEX assesses decision alternatives using multiple criteria. Alternatives are described with variables, called attributes, which represent observable properties of alternatives, such as *Price*, *Availability*, or *Accuracy*. DEX has the following distinctive characteristics:

Table 1 Methods selected for evaluation and comparison within the developed DSS

Method	Detection principle	Reference
qPCR event	Simplex real-time PCR	Holck et al. (2002)
qPCR screen	Simplex real-time PCR	Alary et al. (2002), Kuribara et al. (2002)
qPCR triplex 35S-Iec1-IPC	Multiplex real-time PCR	Pla et al. (2013)
qPCR pentaplex	Multiplex real-time PCR	Huber et al. (2013)
SIMQUANT simplex	Simplex real-time PCR with limiting dilutions	Berdal et al. (2008)
SIMQUANT multiplex	Multiplex real-time PCR with limiting dilutions	Berdal et al. (2008)
ddPCR MON810/hmgA	End-point multiplex droplet digital PCR	Morisset et al. (2013)
ddPCR multiplex per ingredient	End-point multiplex droplet digital PCR	Dobnik et al. (2015)
EAT DualChip	Multiplex PCR coupled with microarray detection	Leimanis et al. (2008), Hamels et al. (2009)
Pentaplex-CGE	Multiplex PCR coupled with capillary gel electrophoresis detection	Nadal et al. (2006)
LAMP screening/detection	Loop-mediated isothermal amplification with end-point visual detection	Chen et al. (2011), Wang et al. (2015)
LAMP-BART	Loop-mediated isothermal amplification with real-time bioluminescence detection	Kiddle et al. (2012)
LFD	Lateral flow device for antibody-based visual detection	Van Den Bulcke et al. (2007)
NGS-flanking regions	Enrichment PCR coupled with next-generation sequencing	Arulandhu et al. (2016)*
NGS-wgs	Whole genome sequencing	Holst-Jensen et al. (2016)*, Willems et al. (2016)

*Review publications with limited experimental data reported

- DEX models are *hierarchical*. Attributes are hierarchically organized, so that the attributes at higher levels of the hierarchy depend on (and are determined on the basis of) lower-level attributes. This effectively splits attributes into basic attributes (terminal nodes), which represent inputs to the model, and aggregated attributes (internal nodes), which represent evaluations of alternatives. The topmost node(s) represent the final evaluations.
- DEX is a *qualitative* MCDA method. Unlike the majority of MCDA methods, which use numerical attributes, DEX uses symbolic attributes. Each attribute in DEX has a value scale consisting of words, such as (no, yes) or (low, medium, high). Scales are usually small (up to five values) and, whenever possible, preferentially ordered from “bad” to increasingly “good” values, according to the purpose of the choice.
- DEX is a *rule-based* method. The evaluation of alternatives is defined in terms of decision rules, which are defined by the model developer (expert of decision maker) and represented in the form of decision table. Each aggregated attribute in the model has a decision table that determines its value for all possible combinations of values of descendant attributes in the hierarchy.

The method DEX is currently implemented in the software called DEXi (Bohanec 2015) (downloadable from <http://kt.ijs.si/MarkoBohanec/dexi.html>). DEXi supports the development of DEX models and their application for the evaluation and analysis of decision alternatives. In the model

development stage, DEXi checks the quality of decision rules in terms of completeness (providing evaluation results for all possible inputs) and consistency (defined decision rules obey the principle of dominance, i.e., they monotonically increase with increasing input values).

DEX and DEXi have so far been used to support real-world decisions in health care, public administration, agriculture, food production, ecology, land use planning, tourism, housing, traffic control, sports, and finance (Bohanec et al. 2013b). Some recent large-scale applications include assessment of food and feed for the presence of genetically modified organisms (Bohanec et al. 2017) and assessment of energy production technologies (Kontić et al. 2016). Overall, DEX is particularly suitable for helping to solve complex decision problems that require judgment and qualitative knowledge-based reasoning, dealing with inaccurate and/or missing data, and analyzing and justifying the results of evaluation (Bohanec et al. 2013b).

MCDA Model for the Assessment of Analytical Methods

The developed *AnalyticalMeth* model (available as *.dxi* file in Online Resource 1) is suitable for evaluation of analytical detection/quantification methods. Its overall architecture is shown in Fig. 1. As input, the model takes data describing analytical methods of the corresponding type. As output, the model provides two assessments: *FitForPurpose*, which tells whether a method is appropriate for a given analytical purpose

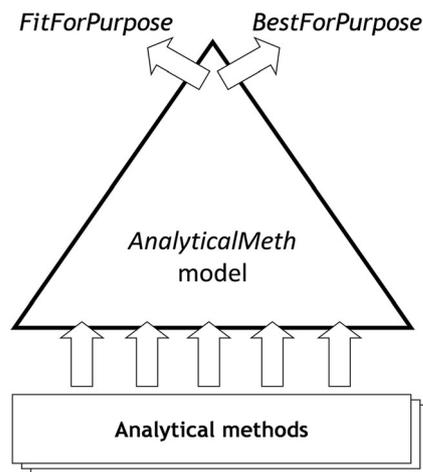


Fig. 1 Schematic structure of the *AnalyticalMeth* model. The two main output attributes (*FitForPurpose* and *BestForPurpose*) appear on the upper side, and the input (basic) attributes that describe the detection and quantification analytical methods appear as terminal nodes at the bottom. Internal (aggregate) attributes serve for the aggregation of basic attributes into the three overall assessments. The aggregation is governed by expert-defined decision rules

(using the value scale (no, partly, yes)), and *BestForPurpose*, which assesses the method's quality, depending on its fitness for purpose, level of development, performance, and overall applicability (using the value scale (unacc, acc, good, v-good, exc)).

The model has a complex internal structure and contains 77 attributes (34 basic, 10 linked, and 33 aggregated). The 34 basic attributes, represented with scales and descriptions, are listed in Table 2. Figure 2 shows the hierarchical structure of attributes in the *AnalyticalMeth* model. Fitness for purpose is assessed through the *FitForPurpose* submodel, which includes two branches: *PurposeFitness* and *SiteFitness*. *PurposeFitness* determines quantitative and screening performance of methods based on the following properties: linearity, accuracy, absolute and relative limit of quantification (LOQ), specificity, robustness, and limit of detection (LOD). The second branch, *SiteFitness*, determines whether the method is fit for on-site applications, meaning that it uses portable and less expensive equipment and that the actual analysis can be performed on the site of sampling (e.g., on the field). The most important part of the model is the *BestForPurpose* branch, which assesses the overall quality of the method in terms of *Constraints* and *Method Evaluation*. The *Constraints* submodel requires that the method is fit for purpose (*FitForPurpose*) and sufficiently developed (*MethDeveloped*) in terms of availability of reference materials (*SuggestedSamples*), defined standard operating procedure (*SOP*), current stage of development (*DevelopmentStage*), known specificity, and proficiency test outcomes. The second submodel, *MethEvaluation*, includes *Costs* (fixed and running costs), *Method Performance* in relation to its primary purpose (detection and/or quantification), and *Method Applicability* (the set of different functionalities of the method, see Fig. 2 and Table 2).

The hierarchical structure of the model provides a framework for decision tables, which define the bottom-up aggregation of attributes in the model. For each of the 33 aggregated attributes, a corresponding decision table was defined. We illustrate here the concept of decision tables with only two examples (Tables 3 and 4); for all decision tables, please see the general model description in Online Resource 2. The decision rules are indeed critical to the final outcome of the evaluation and selection process. If, for example, the laboratory must be able to identify and/or quantify a particular group of targets, then the rules must be designed so that only methods compliant with such a requirement are accepted.

As a first simple example, let us consider the decision table that assesses whether or not an analytic method is suitable for quantification. In the model (Fig. 2), the corresponding aggregated attribute is *FitForQuantification*, which depends on two descendant attributes: *FitForScreening* and *QuantitativePerformance*. The corresponding decision table (Table 3) thus defines the value of *FitForQuantification* for all possible value combinations of the latter attributes. There are nine possible combinations, from

Rule 1: if *FitForScreening* = no and *QuantitativePerformance* = unacc then *FitForQuantification* = no

to

Rule 9: if *FitForScreening* = yes and *QuantitativePerformance* = good then *FitForQuantification* = yes

This decision table is complete (defined for all nine possible combinations) and monotonically increasing.

For a more complex example of an evaluative decision table, let us consider the root attribute *AnalyticalMeth*, which makes an overall assessment of the method combining the attributes *FitForPurpose* (no, partly, yes) and *BestForPurpose* (unacc, acc, good, v-good, exc). Thus, the corresponding decision table contains $3 \times 5 = 15$ combinations. To save space, and in contrast with Table 3, Table 4 shows the rules in a more compact way, employing the symbols “:” (interval), “*” (any value), and “>=” (at least as good as). For example, two of the rules are interpreted as:

Rule 1: if the method is not *FitForPurpose*, then it is unacceptable (regardless of *BestForPurpose*).

Rule 4: if the method is at least partly *FitForPurpose* and is acceptable with respect to *BestForPurpose*, then it is acceptable.

In standard DEX notation for input data that is unknown or so uncertain that it cannot be represented by a single scale value, “*” is used. In our case, the “*” is most often used in connection with *LOQ_Abs*, *LOQ_Rel*, and *Linearity* because these parameters were not relevant for specific methods (i.e., quantification

Table 2 List of all basic attributes and their scales and descriptions

Basic attribute	Scale	Description of scale values	Description of attribute
Purpose	Screening, quantification		Primary purpose of analysis: screening or quantification
Site	Lab, on-site		Site of performing analysis: lab or on-site
DevelopmentStage	Proof, opt. assay, tested, prevalidation, validation		Current state of method development
SOP	No, yes		Is standard operating procedure available?
SuggestedSamples	No, yes		Are samples for method development available? At least reference material and one independent positive sample are in general needed to fully assess the method
SpecificityAssessed	No, partly, yes, not required	No, only for targets available, for all relevant targets, not applicable for the method	To which extent was method specificity assessed? Definition of specificity: property of a method to respond exclusively to the characteristic or analyte of interest
Proficiency	No, yes		Was the method tested on proficiency samples?
MethodType	Detection, quantification		
Specificity	No, yes		Is the method specific?
LOD_Abs	Insensitive, med, good	Above 100, 25–100, 1–25 target copies	Absolute limit of detection: target copy numbers detected in single reactions
LOD_Rel	Insensitive, med, good	Above 0.5% GMO, 0.1–0.5% GMO, below 0.1% GMO	Relative limit of detection: lowest detected percentage of GMO in the sample
Robustness	Low, med, high	The method is running on one machine type only, needs skilled operators, difficult to transfer to other system; the method is well transferable but is sensitive to deviations in experimental conditions; the method is well transferable and not sensitive to experimental conditions	Is the method robust? Definition: the robustness of a method is a measure of its capacity to remain unaffected by small, but deliberate deviations from the experimental conditions described in the procedure. Examples of factors that a robustness test could address are use of different instrument types, operator, brand of reagents, concentration of reagents, and temperature of reaction
InhibitorHandling	None, med, good	Method is sensitive to inhibitors, the method is partially sensitive to inhibitors, the method is not sensitive to inhibitors	How is method handling inhibitor-rich samples?
Accuracy	Low, med, high	> 50% discrepancy, 25–50%, < 25%	How accurate is the method? Definition of accuracy: the closeness of agreement between a test result and the accepted reference value
LOQ_Rel	Bad, med, good	Above 1% GMO, 0.1–1% GMO, below 0.1% GMO	Relative limit of quantification: lowest quantifiable percentage of GMO in sample
LOQ_Abs	Bad, med, good	Above 200, 50–200, below 50 target copies	Absolute limit of quantification: target copy numbers quantified in single reactions
Linearity	Low, med, high	$R^2 < 0.95$ and dynamic range (DR) spanning 2 logs, $R^2 > 0.96$ and DR spanning 2–4 logs, $R^2 > 0.96$ and DR spanning more than 4 logs	How linear is the method within range of quantification (= coefficient of determination, R^2 of linear regression)? What is the DR of the method? Definition of dynamic range: The range of concentrations over which the method performs in a linear manner with an acceptable level of accuracy and precision.
MixedSamples	No, yes		Is the method applicable for mixed samples?
ProcessedSamples	No, yes		Is the method applicable for processed samples?
RequiredAmount	High, med, low	> 5, 1–5, < 1 µg	Required amount of DNA
Screening	No, yes		Is the method useful for screening?
Identification	No, yes		Is the method useful for identification of specific GM line?
SpeciesDetection	No, yes		Is the method useful for species detection?
Quantification	No, yes		Is the method useful for quantification?

Table 2 (continued)

Basic attribute	Scale	Description of scale values	Description of attribute
UnauthorisedGM	No, yes		Can the method detect unauthorized GMOs?
OnSiteDetection	No, yes		Is the method suitable for on-site detection?
Consumables/Sample Samples/Day	High, med, low Low, med, high	Higher than PCR, as in PCR, lower than PCR Under 10, 10–50, more than 50	Costs of consumables per sample compared to simplex PCR Number of samples that can be analyzed in one working day with one detection instrument
Targets/Method	Low, med, high, v-high	1, 2 to 4, 5 to 10, more than 10	Number of targets that can be detected simultaneously
PersonalCosts	High, med, low	More than 3 h, 30 min to 3 h, less than 30 min	Costs of labor per sample [working hours needed per sample] (not including sampling, sample processing, and DNA extraction)
EquipmentCosts	V-high, high, med, low	Range of microarray or ddPCR, range of conv. PCR, basic lab or even on-site Mol. lab, basic lab, no lab needed	Costs of required equipment, including maintenance
InfrastructureCosts	High, med, low	Experienced techn., any techn. or student, not qualified personnel	Cost of infrastructure
CompetenceCosts	High, med, low		Personal competence required for method
ValidationCosts	High, med, low	Range of multiplex methods, range of simplex qPCR, lower	Cost of in-house validation (implementation in the lab)

parameters are not relevant for qualitative methods). Also, the values of *Robustness*, *InhibitorHandling*, and *Accuracy* for the method next-generation sequencing-whole genome sequencing (NGS-wgs) are missing because they were not assessed yet. When evaluating analytical methods with unknown input values, DEX treats the symbol “*” as a set of all possible values of the corresponding attribute and repeats the evaluation for all of them. Consequently, any DEX evaluation may yield a set of values rather than a single value (Bohanec 2015).

The decision tables in the *AnalyticalMeth* model are complete and consistent. All decision tables are presented in the model description (see Online Resource 2).

Results

Filling Up the Model with Methods' Information

The information about the individual methods and their performance must be manually entered into the model. Different variables or observable properties of the methods, called attributes, are taken into account for final evaluation. For each basic attribute, a corresponding scale value must be chosen. For the GMO-related examples presented in this study, the 34 basic attributes and their scale values are listed in Table 2. We determined a specific scale value of each basic attribute for each individual method listed in Table 1. We have performed this task with the help of published data and with the data from our experiments for yet unpublished methods. The constructed methods/attributes matrix (presented in Table 5) is the base for the calculations of the model.

Overall Evaluation of Compared Methods

The developed model was tested using 15 analytical methods. The overall assessment of these 15 methods with the *AnalyticalMeth* model on the basis of the applied decision rules is shown in Fig. 3. Other decision rules might or would have yielded other results. The gold standard technologies, qPCR event and qPCR screen, were assessed as very good and good, respectively. The majority of other methods (qPCR triplex, qPCR pentaplex, SIMQUANT multiplex, both ddPCR methods, EAT Dual Chip, pentaplex-CGE, and LAMP-BART) were assessed as very good and one (SIMQUANT simplex) as good. Only one technology was assessed as excellent, i.e., simplex LAMP for GMO screening or detection. In the current context, the LFD technology was assessed as acceptable, due to weaker performance in terms of sensitivity and accuracy, despite that it can be improved by subsampling strategies (Remund et al. 2001; Kobilinsky and Bertheau 2005). Both NGS methods were assessed as unacceptable, mostly because their sensitivity and throughput are not as good as with other methods and also the price of analyses is relatively

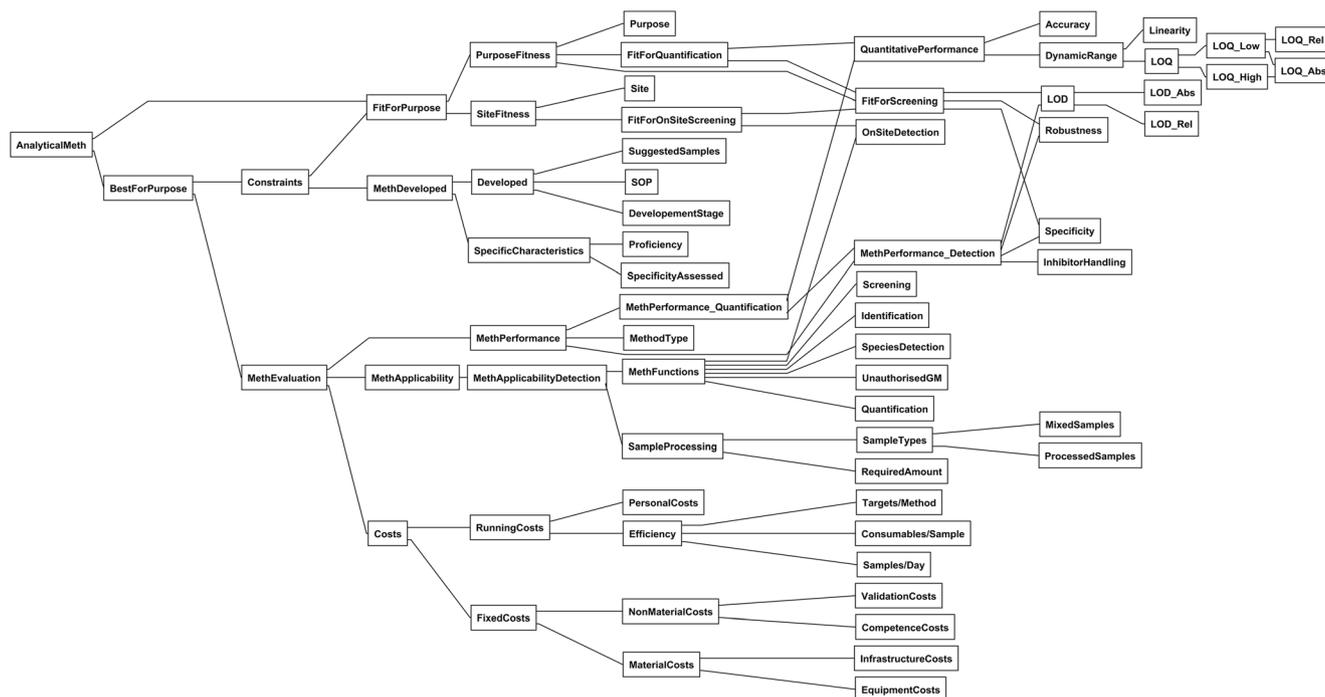


Fig. 2 Hierarchical structure of attributes in the *AnalyticalMeth* model

high. In case of different rules (one example in Table 6), for instance when there would be a need to perform the detection and identification of unauthorized GMOs, the NGS methods would get a better (or even the best) score.

Detailed Method Assessment

To explain the overall assessment and to pinpoint the differences that contribute the most to the outcome of method assessments, a more detailed analysis can be performed when assessing the methods at a lower level. To illustrate this, we have selected three cases to evaluate the effectiveness of increasing the number of targets per analysis (Fig. 4), to compare different detection platforms (Fig. 5) and to compare the methods with the same

purpose (e.g., quantification, Fig. 6). For these cases, we have selected five sublevels, which are, in our opinion, the most informative for the given situation (*BestForPurpose*, *Costs*, *MethPerformance*, *MethApplicability*, and *Constraints* for the first two cases and *QuantitativePerformance*, *Costs*, *MethFunctions*, *LOD*, and *Targets/Method* for quantitative methods comparison). For their position within the attribute tree, see section A in Online Resource 2.

Increasing the Number of Targets per Analysis

In order to compare the performance of qPCR-based detection methods when increasing the number of targets detected in a single analysis, we have included simplex qPCR (event

Table 3 Decision table defining *FitForQuantification* output depending on *FitForScreening* and *QuantitativePerformance* inputs

Rule no.	Input		Output
	<i>FitForScreening</i>	<i>QuantitativePerformance</i>	
1	No	Unacc	No
2	No	Acc	No
3	No	Good	No
4	Partly	Unacc	No
5	Partly	Acc	Partly
6	Partly	Good	Yes
7	Yes	Unacc	No
8	Yes	Acc	Partly
9	Yes	Good	Yes

unacc unacceptable, acc acceptable

Table 4 Decision table for the assessment of *AnalyticalMeth* from *FitForPurpose* and *BestForPurpose* inputs

Rule no.	Input		Output
	FitForPurpose	BestForPurpose	AnalyticalMeth
1	No	*	Unacc
2	*	Unacc	Unacc
3	Partly	Acc:good	Acc
4	>=partly	Acc	Acc
5	Partly	>=v-good	Good
6	Yes	Good	Good
7	Yes	V-good	V-good
8	Yes	Exc	Exc

unacc unacceptable, *acc* acceptable, *v-good* very good, *exc* excellent, “:” interval, “*” any value, “>=” at least as good as

specific and screening), multiplex qPCR (triplex and pentaplex), and SIMQUANT methods. In the overall evaluation, triplex qPCR screen and SIMQUANT simplex were assessed as good, while all the others were assessed as very good (Fig. 3). As can be seen from Fig. 4, lower overall score partially comes from lower scores for the aggregate attributes *BestForPurpose*, *MethApplicability*, and/or *Costs*. When multiplexing and comparing simplex and multiplex of the same platform, more targets are detected in one reaction. Thus, the cost for analysis per target is generally lower. Based on the set decision rules, triplex qPCR was still in the same cost level as both simplex qPCR methods, but pentaplex qPCR already showed cost benefit. The SIMQUANT method not only improved the cost factor with multiplexing, but also the *BestForPurpose* score was better (Fig. 4).

Different Detection Platforms

In order to compare different detection platforms, considering their complete application potential in terms of their purpose, we assessed nine different methods (Fig. 5). We observed that the majority of detection methods alternative to qPCR (droplet digital PCR, microarray detection, capillary gel electrophoresis, and LAMP coupled with bioluminescence detection in real-time) are comparable between each other, i.e., having the same scores, when performing evaluation with the given decision rules. A relatively simple and inexpensive method for on-site detection, LFD, was assessed as less advantageous in our conditions of use, mostly because in general, it targets proteins and thus cannot be applied to the analysis of processed samples (Fig. 5). The NGS method is currently, based on the scores of the *AnalyticalMeth* model, still not a good option for GMO analysis. Nevertheless, NGS is the newest fashionable method, and with more extensive use driven by a predicted drop in overall costs it might find a place in GMO testing in the future, when sequencing methods will have been improved for in-

depth sequencing, sequences' assembly, and comparison with, currently missing, gold reference genomes. Already now, it is the best option for very specific applications when no other methods are applicable (e.g., for identification of completely unknown GMO), and with new platforms it is getting less expensive (Pennisi 2017). Another thing is that while new alternative detection platforms often have very good applicability, the fixed costs (equipment, training, skilled personnel, etc.) for implementation in the laboratory can be prohibitively high, resulting in a low overall score in this *AnalyticalMeth* model (e.g., NGS in Fig. 5).

Methods for Target Quantification

The comparison of the methods that enable quantification of the targets, including the current golden standard detection method (qPCR on transformation event's sequences), is shown in Fig. 6. For this comparison, we selected the attributes that are important for quantification methods and show most pronounced differences. For example, the attribute *FitForQuantification* would not provide any new relevant information, as all of the methods except NGS-wgs have reported quantification purposes. However, *MethFunctions*, and *Costs*, with sublevel of *Targets/Method*, produce a lot of information, if one is planning to implement the method into routine diagnostics. Additionally, *QuantitativePerformance* and *LOD* (for aspect of sensitivity) attributes were added to the comparison. With such a selection of attributes, the output results of the *AnalyticalMeth* model were the most relevant for the purpose of GMO quantification. The comparison of SIMQUANT simplex and multiplex shows improvement in terms of costs (with more targets analyzed per method). However, when moving to ddPCR (considering it as a technical improvement of the SIMQUANT method), we also gain additional information within *MethFunctions* due to simultaneous species reference sequence and event detection (Fig. 6). When finally moving to ddPCR multiplex per ingredient, we got the best scores in this comparison with the highest number of targets. When comparing ddPCR multiplex to a qPCR event method, ddPCR multiplex outperforms qPCR event due to the possibility of endogenous sequence quantification (and species identification) and due to numerous targets quantified simultaneously in one reaction (Fig. 6). NGS-wgs that theoretically also enables quantification, has its strong points in detecting/quantifying numerous targets (events, species, and also unauthorized GMOs). However, for the current case study of a routine laboratory detecting approved GMOs, other attributes of NGS-wgs presented low scores; therefore, the overall score for this method is relatively poor in comparison to other methods (Fig. 6). In case that a model would be designed to evaluate the best method for the purpose of unauthorized GM detection, the overall score of NGS-wgs would probably be the highest.

Table 5 Methods and attributes matrix in the model. For individual method a value is selected for each of the attributes

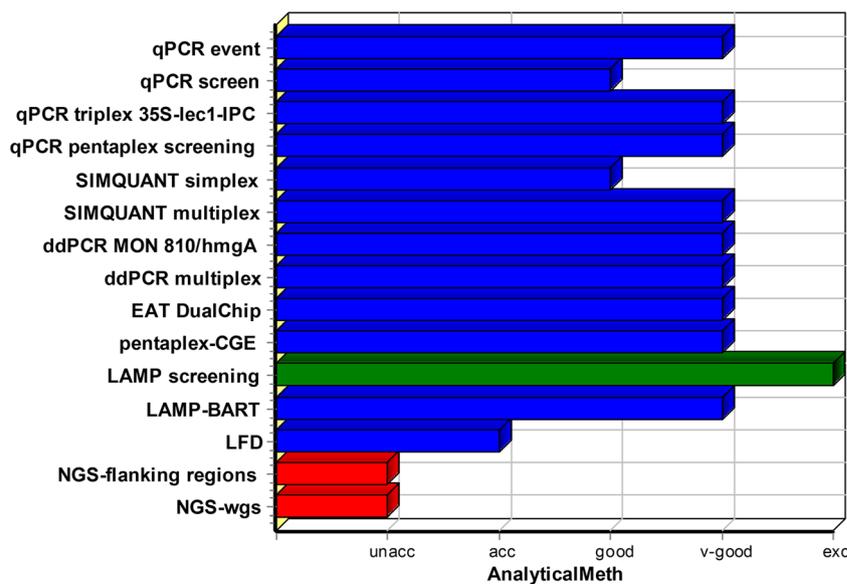
Attribute	qPCR event	qPCR screen	qPCR triplex 35S-lecl-IPC	qPCR pentaplex screening	SIMQUANT simplex	SIMQUANT multiplex	ddPCR MON 810/hmgA
Reference	Holick et al. (2002)	Alary et al. (2002), Kuribara et al. (2002)	Pla et al. (2013)	Huber et al. (2013)	Berdal et al. (2008)	Berdal et al. (2008)	Morisset et al. (2013)
Purpose	Quantification	Screening	Screening	Screening	Quantification	Quantification	Quantification
Site	Lab	Lab	Lab	Lab	Lab	Lab	Lab
DevelopmentStage	Validation	Validation	Validation	Validation	Validation	Validation	Validation
SOP	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SuggestedSamples	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SpecificityAssessed	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Proficiency	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MethodType	Quantification	Detection	Detection	Detection	Quantification	Quantification	Quantification
Specificity	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LOD_Abs	Good	Good	Good	Good	Good	Good	Good
LOD_Rel	Good	Good	Good	Good	Good	Good	Good
Robustness	High	High	High	High	High	High	High
InhibitorHandling	Good	Good	Good	Good	Good	Good	Good
Accuracy	High	High	High	High	High	High	High
LOQ_Rel	Good	*	*	*	Good	Good	Good
LOQ_Abs	Good	*	*	*	Good	Good	Good
Linearity	High	High	High	High	High	High	High
MixedSamples	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ProcessedSamples	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RequiredAmount	Low	Low	Low	Low	Low	Low	Low
Screening	No	Yes	Yes	Yes	Yes	Yes	No
Identification	Yes	No	No	No	Yes	Yes	Yes
SpeciesDetection	No	No	Yes	No	No	No	Yes
Quantification	Yes	No	No	No	Yes	Yes	Yes
UnauthorisedGM	No	No	No	No	No	No	No
OnSiteDetection	No	No	No	No	No	No	No
Consumables/Sample	Med	Med	Med	Med	High	High	Med
Samples/Day	High	High	High	High	High	High	Med
Targets/Method	Low	Low	Med	High	Low	Med	Med
PersonalCosts	Med	Med	Med	Med	Med	Med	Med
EquipmentCosts	High	High	High	High	High	High	V-high
InfrastructureCosts	High	High	High	High	High	High	High
CompetenceCosts	Med	Med	Med	Med	Med	Med	High
ValidationCosts	Med	Med	High	High	Med	High	High
Attribute	ddPCR multiplex per ingredient	EAT DualChip	Pentaplex-CGE	LAMP screening/detection	LAMP-BART	LFD	NGS-flanking regions NGS-wgs
Reference	Dobnik et al. (2015)	Leimanis et al. (2008), Hamels et al. (2009)	Nadal et al. (2006)	Chen et al. (2011), Wang et al. (2015)	Kiddle et al. (2012)	Van Den Bulcke et al. (2007)	NA
Purpose	Quantification	Screening	Screening	Screening	Screening	Screening	Quantification

Table 5 (continued)

Attribute	ddPCR multiplex per ingredient	EAT DualChip	Pentaplex-CGE	LAMP screening/detection	LAMP-BART	LFD	NGS-flanking regions	NGS-wgs
Site	Lab	Lab	Lab	On-site	On-site	On-site	Lab	Lab
Development/Stage	Validation	Validation	Validation	Validation	Validation	Validation	Proof	Proof
SOP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
SuggestedSamples	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Specificity/Assessed	Yes	Yes	Yes	Yes	Yes	Yes	Partly	Partly
Proficiency	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Method/Type	Quantification	Detection	Detection	Detection	Detection	Detection	Detection	Quantification
Specificity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LOD_Abs	Good	Med	Good	Good	Good	Med	Med	Med
LOD_Rel	Good	Med	Good	*	Good	Med	Insensitive	Insensitive
Robustness	High	High	High	High	High	Med	Med	*
Inhibitor/Handling	Good	Good	Good	Good	Good	Good	Med	*
Accuracy	High	High	High	High	High	Med	Med	*
LOQ_Rel	Good	*	*	*	*	*	*	Bad
LOQ_Abs	Good	*	*	*	*	*	*	Bad
Linearity	High	Med	High	*	High	*	*	*
MixedSamples	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ProcessedSamples	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
RequiredAmount	Low	Med	Low	Low	Low	Low	Low	Med
Screening	No	Yes	Yes	Yes	Yes	No	Yes	Yes
Identification	Yes	Yes	Yes	No	No	Yes	Yes	Yes
SpeciesDetection	Yes	Yes	Yes	No	No	No	Yes	Yes
Quantification	Yes	No	No	No	No	No	No	Yes
UnauthorisedGM	No	No	No	No	No	No	Yes	Yes
OnSiteDetection	No	No	No	Yes	Yes	Yes	No	No
Consumables/Sample	Med	Low	Low	Med	Med	Low	High	High
Samples/Day	Med	Low	Med	High	High	High	Low	Low
Targets/Method	V-high	V-high	High	Low	Low	Low	V-high	V-high
PersonalCosts	Med	High	Med	Low	Med	Low	High	High
EquipmentCosts	V-high	V-high	V-high	Low	Med	Low	V-high	V-high
InfrastructureCosts	High	High	High	Low	Low	Low	High	High
CompetenceCosts	High	High	High	Low	Med	Low	High	High
ValidationCosts	High	Med	High	Med	Med	Med	High	High

NA not available, because the method was not published yet; *unacc* unacceptable; *acc* acceptable; *v-high* very high; “*” any value

Fig. 3 Overall evaluation of the methods by the *AnalyticalMeth* model given the specific set of decision rules defined for the purpose of GMO detection or quantification in Online Resource 2 part B. Colors of the chart are for easier visualization of the results, as they visually stress the appropriateness of the methods (green—excellent result; blue—acceptable to very good result; red—unacceptable result)



Discussion

The presented results from the *AnalyticalMeth* model showed its usefulness for evaluating a set of GMO testing methods of diverse detection platforms. The results also showed that new

methods, developed for the purpose of GMO detection, are good, or even better, when compared to the golden standard method of qPCR. Most of the compared methods showed a very good final evaluation score. However, it should be noted that the comparison of scores at sublevels varied between the

Table 6 Decision table for the assessment of *MethFunctions* with theoretical example of assessment output, when the laboratory would require detection and identification of unauthorized GMO

Rule no.	Input						Output	
	Screening	Identification	SpeciesDetection	Quantification	UnauthorisedGM	OnSiteDetection	MethFunctions	MethFunctions—alternative ^a
1	No	No	No	*	*	*	Bad	Bad
2	No	No	Yes	No	No	No	Acc	Bad
3	No	Yes	No	No	No	No	Acc	Acc
4	Yes	No	No	No	No	No	Acc	Bad
5	No	No	Yes	Yes	No	No	Good	Bad
6	*	Yes	No	Yes	No	No	Good	Good
7	*	Yes	Yes	No	No	No	Good	Good
8	Yes	*	No	Yes	No	No	Good	Bad
9	Yes	*	Yes	No	No	No	Good	Bad
10	Yes	Yes	No	*	No	No	Good	Good
11	Yes	Yes	*	No	No	No	Good	Good
12	*	*	Yes	*	*	Yes	Exc	Bad
13	*	*	Yes	*	Yes	*	Exc	Exc
14	*	Yes	*	*	*	Yes	Exc	Good
15	*	Yes	*	*	Yes	*	Exc	Exc
16	*	Yes	Yes	Yes	*	*	Exc	Good
17	Yes	*	*	*	*	Yes	Exc	Bad
18	Yes	*	*	*	Yes	*	Exc	Exc
19	Yes	*	Yes	Yes	*	*	Exc	Bad

^a The suggested rule would favor the detection and identification of unauthorized GMOs

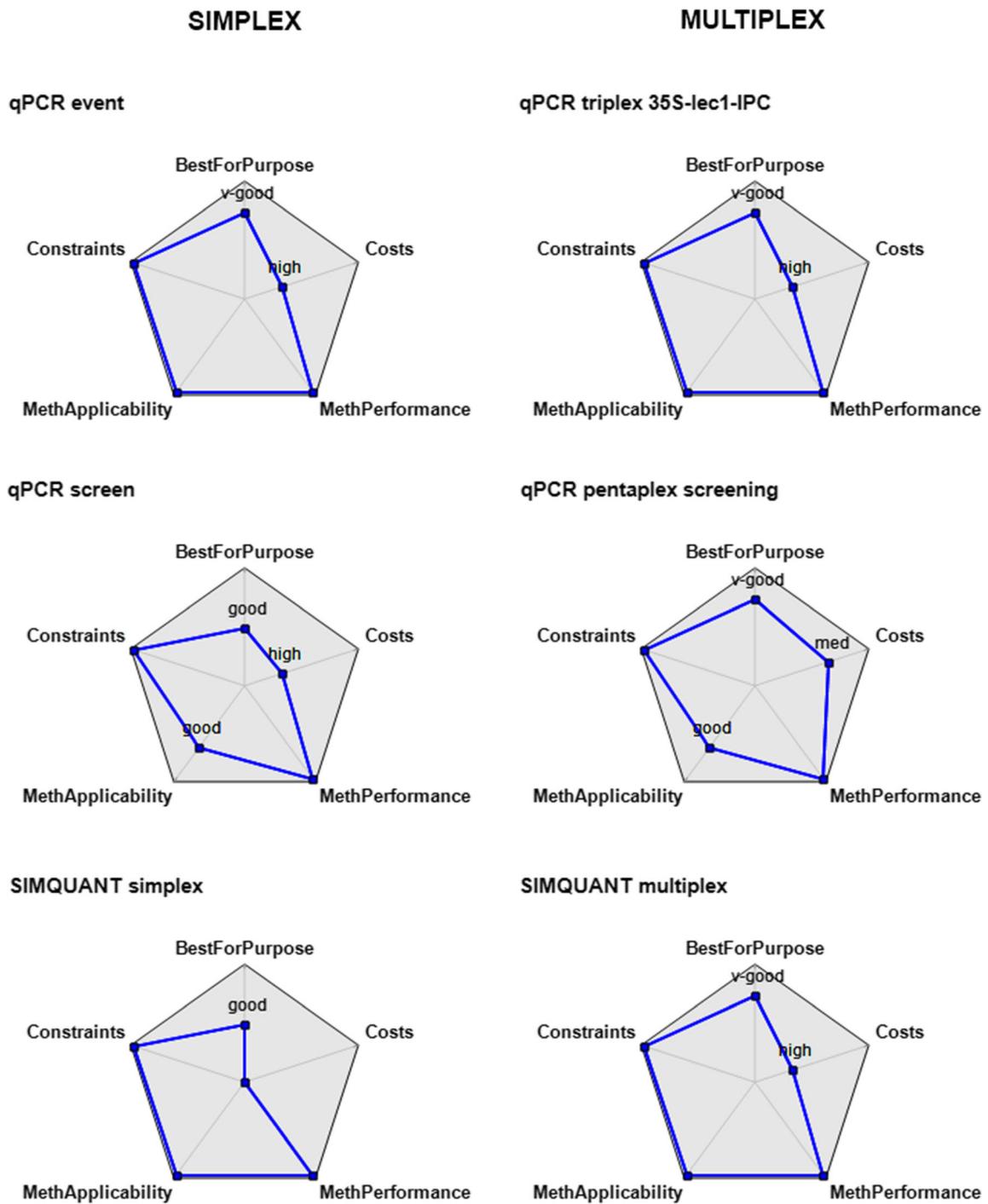


Fig. 4 Comparison of qPCR-based methods to evaluate the effectiveness of increasing the number of targets per analysis. The left side of the figure shows evaluation of simplex methods with qPCR event, qPCR screen, and SIMQUANT simplex as examples. On the right side, a multiplex method evaluation is given for qPCR triplex, qPCR pentaplex, and

SIMQUANT multiplex. Both qPCR pentaplex and SIMQUANT multiplex showed improvement in costs, whereas triplex qPCR was still in the range of simplex qPCR. Blue line—overall score for evaluation of the method was “acceptable,” good, or very good

methods. As the compared methods are meant for different purposes, the comparison at these sublevels is indeed more informative for the laboratories. Importantly, the final evaluation score depends on the decision rules set by the user. Therefore, the current model can only serve as an example

for the decision rules set by ourselves. Indeed, the decision rules, giving more importance to other attributes, set by other users, might result in different evaluation scores in the end.

For our model, one of the important attributes was the cost effectiveness of a method. Since the costs in this model are

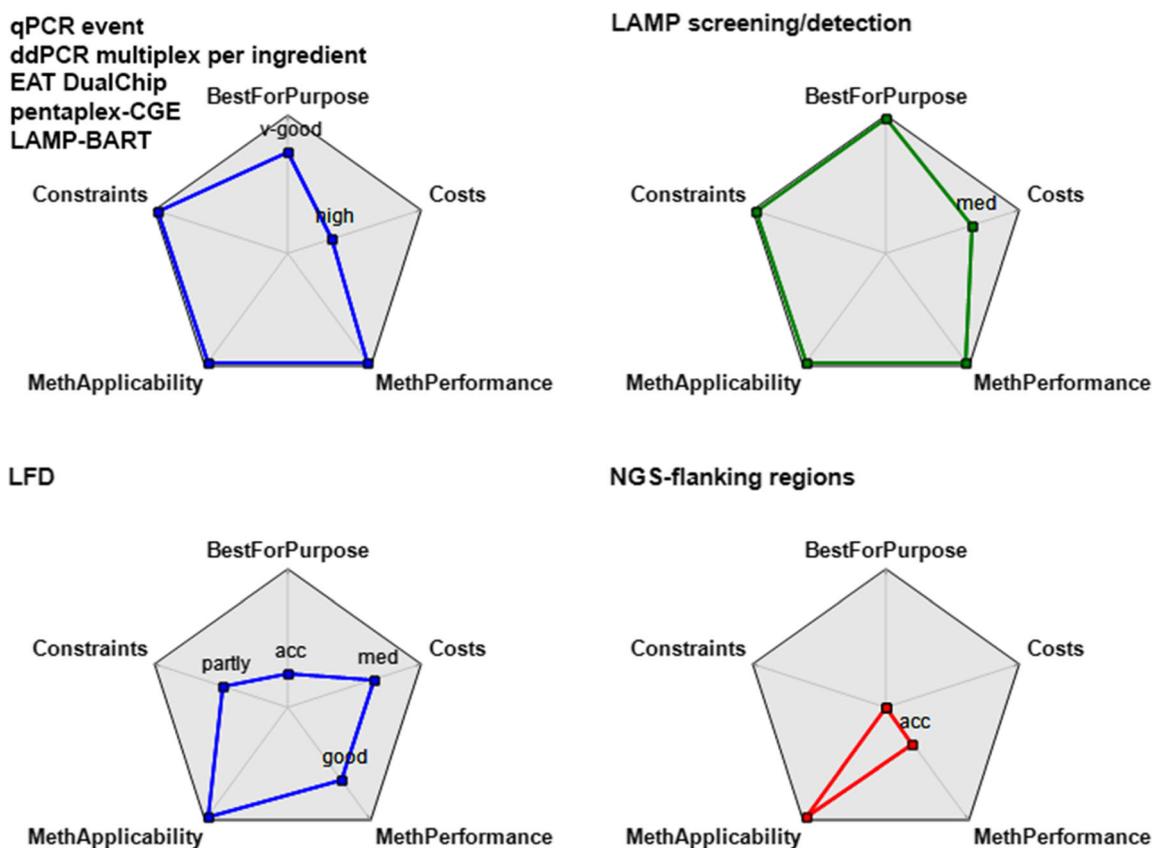


Fig. 5 Comparison of different detection platforms. Majority of platforms have equal score in terms of selected attributes (top left) and three other platforms have specific scores that are better (LAMP platform) or worse than majority (LFD and NGS). Green line—overall score for

evaluation of the platform was “excellent.” Blue line—overall score for evaluation of the platform was “acceptable,” good, or very good. Red line—overall score for evaluation of the platform was “unacceptable”

compound of running and fixed costs, each one with additional sublevels, the final score of different methods could be comparable. In such cases, one should carefully evaluate also other levels to take a decision, which method should be implemented in the laboratory. If fixed costs are low, but running costs are high, then laboratories with lots of samples might rather make a choice of methods with higher fixed costs and lower running costs as it would be more cost-effective in the long run.

New multiplexing methods (e.g., multiplex ddPCR) do show some increased performance in comparison to simplex qPCR methods, as the number of targets detected in one analysis far exceeds the single target from simplex methods (Morisset et al. 2013; Dobnik et al. 2015, 2016; Košir et al. 2017). However, their main drawback is the investment in new equipment, additional personnel training costs, and the longer time for analysis of one sample. This is generally a problem of new technologies, for which the costs are quite high compared to already established technologies. On the other hand, they can offer more information from one run, outweighing some additional costs. Pentaplex qPCR (Huber et al. 2013) exemplifies this as it incorporates both qPCR and multiplexing, but the purpose

of this method is limited to screening. With the increasing number of GMOs on the market (James 2015), the situation will most probably go in the direction of bringing multiplex methods to the position as *BestForPurpose* technology on the market. Since the quantitative aspect of qPCR multiplex is rather limited, with only two available interlaboratory-validated duplex methods (Waiblinger et al. 2007; Takabatake et al. 2011), new technologies such as ddPCR can become the leading technology on the market for routine GMO diagnostics. When new technologies emerge, they are often more costly at the beginning, but with gradually broader adoption, the prices usually decrease. In the long run, the accumulated costs of delayed implementation can sometimes exceed the accumulated savings perceived in a short-term perspective. This perspective is not included in the cost calculation in the current model, in part because it would add a factor of speculation, as future cost fluctuations cannot be reliably predicted. Nevertheless, up to now in some cases, when considering the need of accurate results for taking some decisions, such as removing products from the market, the current need for parallel use of different NGS platforms and software, to discard tool-linked specific errors,

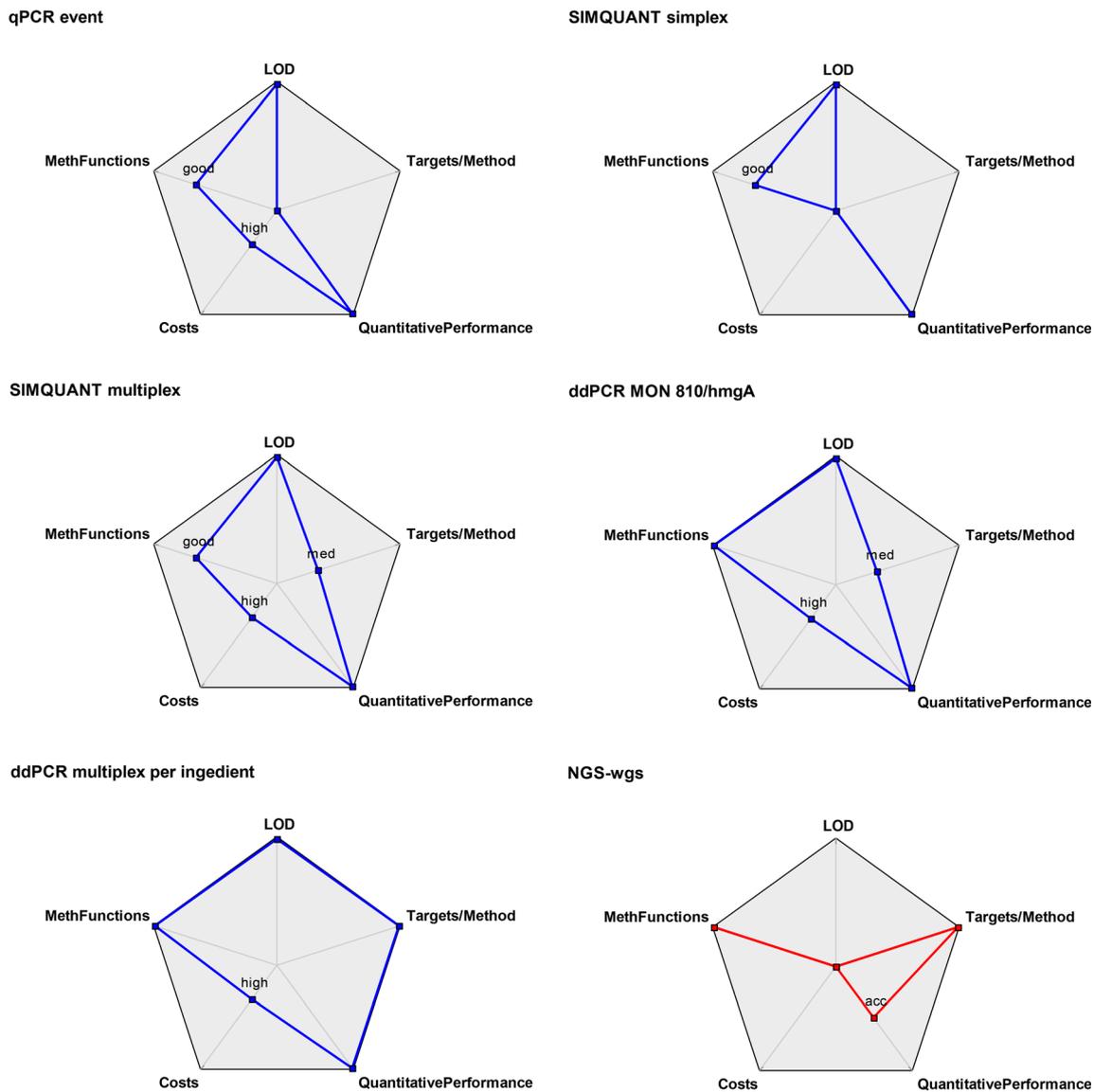


Fig. 6 Comparison of methods that enable GMO quantification. The selected attributes were the ones where the distinction between the methods was the greatest. In comparison to qPCR, only ddPCR showed

improvement in more than two out of five attributes. Blue line—overall score for evaluation of the method was “acceptable,” good, or very good. Red line—overall score for evaluation of the method was “unacceptable”

increases drastically the associated costs, although the costs of NGS are reducing (Liu et al. 2012; Goldfeder et al. 2016; Potapov and Ong 2017).

The GMO analysis testing pipeline generally involves several steps, selected based on a classification by the matrix approach (Chaouachi et al. 2008; Van den Bulcke et al. 2010; Block et al. 2013) and dependent on the sample type. The sample can be analyzed with a multiplex screening method and then further analyzed with a quantitative method. Thus, the comparison of these two methods would not give any relevant information as both are needed in the analysis process. Therefore, we suggest that methods with similar purpose, suitable for each of the steps, should be evaluated independently. Until now, unless a specific GMO is targeted (e.g.,

during an emergency period linked to a specific unauthorized GMO, such as FP967 flax (EURL-GMFF 2009) or BT10 maize (EURL-GMFF 2005)), initial screening is the predominant first step in GMO testing. Based on screening results, it is possible to predict which GM events the sample contains. To eliminate the need for screening steps, a ddPCR multiplex method that quantifies a whole group of GMOs (Dobnik et al. 2015) can for instance be used when only one species is targeted. With the recent emergence and growing number of GMOs lacking the most common screening elements, the cost efficiency of element screening is reduced, since additional complementary identification methods must always be run. In such cases, new methods and more universal approaches could be more suitable for the analysis (e.g., performing

specific event detections with the usual screenings, or using multiplex ddPCR or NGS). Modifications, i.e., setting new decision rules, in the developed MCDA *AnalyticalMeth* model, could help in direct comparison of selected methods.

Again, it is important to note that to get the best possible evaluation and avoid cases where methods would have the same result, the laboratories should define their needs first, and then set up the decision rules accordingly. As two different sets of decision rules might give two different scores for any individual method, the setting of decision rules to fit a laboratory's needs is critical. The developed *AnalyticalMeth* model is therefore not fixed, but fully flexible, to allow each user to select decision rules according to own needs. Additional methods can also be added to the model as they become available. With the emergence of new methods and other relevant parameters, new attributes can easily be added and/or deleted from the model. For instance, seed quality control during production would probably benefit from emphasis on fast on-site applicability. This could improve the ranking of, e.g., LFD, and lower that of PCR. On the other hand, perceived risk of presence of multiple events including unauthorized and possibly even unknown events would suggest to put emphasis on criteria that could favor NGS. One of the possible additions to the model could be a module on DNA extraction, because different methods may require different quantities and purities of DNA. As at least simplified DNA extraction must be performed for methods such as LAMP and by implementing this in the model, its overall on-site applicability might be a bit lower. On the other hand, with available small portable qPCR machines, simplified DNA extraction protocols, and inhibitor-tolerant enzymes for qPCR, a score for qPCR might be a bit higher for on-site detection.

There is actually no limit for the complexity of attributes and rules, which could also include some more laboratory-based observations, such as trust in the reagents (e.g., variability between batches), number and reliability of available reference methods, and amount of costly training needed. Individual laboratories might even put more weight to specific attributes when selecting a series of methods. Such additions to the model could thus provide a substantial contribution to the final evaluation. Since this manuscript compared only the methods that are already publicly available, it is really important that the model offers the possibility of modifying the attributes with emergence of new methods and requirements.

The *AnalyticalMeth* model as presented here has some clear limitations. It can inform but not conclude on which methods to combine for specific aims of GMO testing (e.g., detection only, identification, quantification or detection of unauthorized GMOs). It provides information on a general comparison of individual methods based on their purpose in separate steps of GMO analysis. But it remains open and flexible for future changes that could also set the premises for such even more complex evaluations. The *AnalyticalMeth* model

file is available as Online Resource 1 and can be opened, viewed, and changed using DEXi (downloadable from <http://kt.ijs.si/MarkoBohanec/dexi.html>).

To take one step further, beyond current GMO detection, MCDA could, for instance, consider including epigenomic as well as epitranscriptomic detection by, e.g., sequencing, to be combined with genetic data for, e.g., detecting plants produced with new breeding techniques. Values could be defined for stakeholders who are looking only to detect the products and identification sets for enforcement laboratories, which could be interested in identifying the patent owner of a product or the genome or epigenome modification (i.e., modifications of DNA, associated proteins, and/or RNA) technique used.

Conclusions

The idea of the developed MCDA model was to integrate the evaluation of different GMO detection methods in a decision support system that is operational and easily accessible for various categories of users and that provides data and advice for decision problems that occur in supply chains involving GMOs. In principle, the models' objective was to provide a tool to assess "decision alternatives," to change decision-related parameters and investigate their effects, to visualize the results of evaluations and analyses, and to maintain data related to the decisions involved. We have shown that use of the model can objectively evaluate different kinds of methods that can help in selecting the best for the purpose of interest. Due to the adaptability of the models' generic structure, it can be easily modified for evaluation of methods in other fields.

Acknowledgements We thank Jeroen van Dijk and Bjørn Spilberg for sharing information on unpublished NGS-based methods.

Funding The research leading to these results has received funding from the European Union under FP7 grant agreement no. 613908 (project DECATHLON) and FP6 contract no. 007158 (Project Co-Extra). Complementary financial support was given by the Slovenian Research Agency (contract number P4-0165) and the Research Council of Norway.

Compliance with Ethical Standards

Conflict of Interest David Dobnik declares that he has no conflict of interest. Kristina Gruden declares that she has no conflict of interest. Jana Žel declares that she has no conflict of interest. Yves Bertheau declares that he has no conflict of interest. Arne Holst-Jensen declares that he has no conflict of interest. Marko Bohanec declares that he has no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent Not applicable.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Acutis M, Trevisiol P, Confalonieri R, Bellocchi G, Grazioli E, van den Eede G, Paoletti C (2007) Analytical method performance evaluation (AMPE)—a software tool for analytical method validation. *J AOAC Int* 90:1432–1438
- Alary R, Serin A, Maury D, Ben Jouira H, Sirven JP, Gautier MF, Joudrier P (2002) Comparison of simplex and duplex real-time PCR for the quantification of GMO in maize and soybean. *Food Control* 13: 235–244. [https://doi.org/10.1016/S0956-7135\(02\)00015-4](https://doi.org/10.1016/S0956-7135(02)00015-4)
- Arulandhu AJ, van Dijk JP, Dobnik D, Holst-Jensen A, Shi J, Zel J, Kok EJ (2016) DNA enrichment approaches to identify unauthorized genetically modified organisms (GMOs). *Anal Bioanal Chem* 408: 4575–4593. <https://doi.org/10.1007/s00216-016-9513-0>
- Bahrdt C, Krech AB, Wurz A, Wulff D (2010) Validation of a newly developed hexaplex real-time PCR assay for screening for presence of GMOs in food, feed and seed. *Anal Bioanal Chem* 396:2103–2112. <https://doi.org/10.1007/s00216-009-3380-x>
- Berdal KG, Bøydler C, Tengs T, Holst-Jensen A (2008) A statistical approach for evaluation of PCR results to improve the practical limit of quantification (LOQ) of GMO analyses (SIMQUANT). *Eur Food Res Technol* 227:1149–1157. <https://doi.org/10.1007/s00217-008-0830-1>
- Bertheau Y (2013) Genetically modified and non-genetically modified food supply chains: co-existence and traceability. Wiley-Blackwell, Oxford
- Block A, Debode F, Grohmann L, Hulin J, Tavemiers I, Kluga L, Barbaud-Piednoir E, Broeders S, Huber I, van den Bulcke M, Heinze P, Berben G, Busch U, Roosens N, Janssen E, Žel J, Gruđen K, Morisset D (2013) The GMOseek matrix: a decision support tool for optimizing the detection of genetically modified plants. *BMC Bioinformatics* 14:256. <https://doi.org/10.1186/1471-2105-14-256>
- Bohanec M (2015) DEXi: program for multi-attribute decision making, user's manual, version 5.00. IJS report DP-11897. Ljubljana
- Bohanec M, Bertheau Y, Brera C, Gruđen K, Holst-Jensen A, Kok EJ, Lécroart B, Messéan A, Miraglia M, Onori R, Prins TW, Soler LG, Žnidaršič M (2013a) The co-extra decision support system: a model-based integration of project results. In: Genetically modified and non-genetically modified food supply chains: co-existence and traceability. Wiley-Blackwell, Oxford, pp 459–489. <https://doi.org/10.1002/9781118373781.ch25>
- Bohanec M, Žnidaršič M, Rajković V et al (2013b) DEX methodology: three decades of qualitative multi-attribute modeling. *Inform* 37:49–54
- Bohanec M, Boshkoska BM, Prins TW, Kok EJ (2017) SIGMO: a decision support system for identification of genetically modified food or feed products. *Food Control* 71:168–177. <https://doi.org/10.1016/j.foodcont.2016.06.032>
- Chaouachi M, Chupeau G, Berard A et al (2008) A high-throughput multiplex method adapted for GMO detection. *J Agric Food Chem* 56:11596–11606. <https://doi.org/10.1021/jf801482r>
- Chen L, Guo J, Wang Q, Kai G, Yang L (2011) Development of the visual loop-mediated isothermal amplification assays for seven genetically modified maize events and their application in practical samples analysis. *J Agric Food Chem* 59:5914–5918. <https://doi.org/10.1021/jf200459s>
- Dobnik D, Spilsberg B, Bogožalec Košir A, Holst-Jensen A, Žel J (2015) Multiplex quantification of 12 European Union authorized genetically modified maize lines with droplet digital polymerase chain reaction. *Anal Chem* 87:8218–8226. <https://doi.org/10.1021/acs.analchem.5b01208>
- Dobnik D, Štebih D, Blejec A, Morisset D, Žel J (2016) Multiplex quantification of four DNA targets in one reaction with Bio-Rad droplet digital PCR system for GMO detection. *Sci Rep* 6. <https://doi.org/10.1038/srep35451>
- EURL-GMFF (2005) Detection method for event Bt 10 using a qualitative PCR assay. <http://gmo-crl.jrc.ec.europa.eu/summaries/Bt10.DetectionProtocol.pdf>. Accessed 24 May 2017
- EURL-GMFF (2009) NOST-Spec construct-specific method for the detection of CDC Triffid Flax (Event FP967) using real-time PCR. <http://gmo-crl.jrc.ec.europa.eu/doc/Flax-CDCTriffi>. Accessed 24 May 2017
- European Commission (2003) Regulation (EC) No 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed. *Off J Eur Union L* 1–23
- Germini A, Zanetti A, Salati C, Rossi S, Forré C, Schmid S, Marchelli R, Fogher C (2004) Development of a seven-target multiplex PCR for the simultaneous detection of transgenic soybean and maize in feeds and foods. *J Agric Food Chem* 52:3275–3280. <https://doi.org/10.1021/jf035052x>
- Goldfeder RL, Priest JR, Zook JM, Grove ME, Waggott D, Wheeler MT, Salit M, Ashley EA (2016) Medical implications of technical accuracy in genome sequencing. *Genome Med* 8:24. <https://doi.org/10.1186/s13073-016-0269-0>
- Greco S, Ehrogott M, Figueira J (2016) Multiple criteria decision analysis. Springer, New York. <https://doi.org/10.1007/978-1-4939-3094-4>
- Gruère GP, Rao SR (2007) A review of international labeling policies of genetically modified food to evaluate India's proposed rule 10:51–64
- Hamelis S, Glouden T, Gillard K, Mazzara M, Debode F, Foti N, Sneyers M, Esteve Nuez T, Pla M, Berben G, Moens W, Bertheau Y, Audéon C, van den Eede G, Remacle J (2009) A PCR-microarray method for the screening of genetically modified organisms. *Eur Food Res Technol* 228:531–541. <https://doi.org/10.1007/s00217-008-0960-5>
- Holck A, Vaitilingom M, Didierjean L, Rudi K (2002) 5'-Nuclease PCR for quantitative event-specific detection of the genetically modified Mon810 MaisGard maize. *Eur Food Res Technol* 214:449–454. <https://doi.org/10.1007/s00217-001-0473-y>
- Holck A, Pedersen BO, Heir E (2010) Detection of five novel GMO maize events by qualitative, multiplex PCR and fluorescence capillary gel electrophoresis. *Eur Food Res Technol* 231:475–483. <https://doi.org/10.1007/s00217-010-1302-y>
- Holst-Jensen A (2009) Testing for genetically modified organisms (GMOs): past, present and future perspectives. *Biotechnol Adv* 27: 1071–1082. <https://doi.org/10.1016/j.biotechadv.2009.05.025>
- Holst-Jensen A, Bertheau Y, de Loose M, Grohmann L, Hamelis S, Hougs L, Morisset D, Pecoraro S, Pla M, den Bulcke MV, Wulff D (2012) Detecting un-authorized genetically modified organisms (GMOs) and derived materials. *Biotechnol Adv* 30:1318–1335. <https://doi.org/10.1016/j.biotechadv.2012.01.024>
- Holst-Jensen A, Spilsberg B, Arulandhu AJ, Kok E, Shi J, Zel J (2016) Application of whole genome shotgun sequencing for detection and characterization of genetically modified organisms and derived products. *Anal Bioanal Chem* 408:4595–4614. <https://doi.org/10.1007/s00216-016-9549-1>
- Huber I, Block A, Sebah D, Debode F, Morisset D, Grohmann L, Berben G, Štebih D, Milavec M, Žel J, Busch U (2013) Development and validation of duplex, triplex, and pentaplex real-time PCR screening assays for the detection of genetically modified organisms in food and feed. *J Agric Food Chem* 61:10293–10301. <https://doi.org/10.1021/jf402448y>
- Ishizaka A, Nemery P (2013) Multi-criteria decision analysis: methods and software. Wiley, Somerset
- James C (2015) Global status of commercialized biotech/GM crops: 2015. ISAAA brief no. 51. Ithaca, NY

- Jasbeer K, Son R, Mohamad Ghazali F, Cheah YK (2009) Real-time PCR evaluation of seven DNA extraction methods for the purpose of GMO analysis. *Int Food Res J* 16:329–341
- Kiddle G, Hardinge P, Buttigieg N, Gandelman O, Pereira C, McElgunn CJ, Rizzoli M, Jackson R, Appleton N, Moore C, Tisi LC, Murray JAH (2012) GMO detection using a bioluminescent real time reporter (BART) of loop mediated isothermal amplification (LAMP) suitable for field use. *BMC Biotechnol* 12:15. <https://doi.org/10.1186/1472-6750-12-15>
- Kobilinsky A, Bertheau Y (2005) Minimum cost acceptance sampling plans for grain control, with application to GMO detection. *Chemom Intell Lab Syst* 75:189–200. <https://doi.org/10.1016/j.chemolab.2004.07.005>
- Kontić B, Bohanec M, Kontić D, Trdin N, Matko M (2016) Improving appraisal of sustainability of energy options—a view from Slovenia. *Energy Policy* 90:154–171. <https://doi.org/10.1016/j.enpol.2015.12.022>
- Košir AB, Spilsberg B, Holst-Jensen A, Žel J, Dobnik D (2017) Development and inter-laboratory assessment of droplet digital PCR assays for multiplex quantification of 15 genetically modified soybean lines. *Sci Rep* 7:8601. <https://doi.org/10.1038/s41598-017-09377-w>
- Kuribara H, Shindo Y, Matsuoka T, Takubo K, Futo S, Aoki N, Hirao T, Akiyama H, Goda Y, Toyoda M, Hino A (2002) Novel reference molecules for quantitation of genetically modified maize and soybean. *J AOAC Int* 85:1077–1089
- Leimanis S, Hamels S, Nazé F, Mbongolo Mbella G, Sneyers M, Hohegger R, Broll H, Roth L, Dallmann K, Micsinai A, la Paz JL, Pla M, Brünen-Nieweler C, Papazova N, Taverniers I, Hess N, Kirschneit B, Bertheau Y, Audeon C, Laval V, Busch U, Pecoraro S, Neumann K, Rösel S, van Dijk J, Kok E, Bellocchi G, Foti N, Mazzara M, Moens W, Remacle J, van den Eede G (2008) Validation of the performance of a GMO multiplex screening assay based on microarray detection. *Eur Food Res Technol* 227:1621–1632. <https://doi.org/10.1007/s00217-008-0886-y>
- Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M (2012) Comparison of next-generation sequencing systems. *J Biomed Biotechnol* 2012:1–11. <https://doi.org/10.1155/2012/251364>
- Luan FX, Tao R, Xu YG, Wu J, Guan XJ (2012) High-throughput detection of genetically modified rice ingredients in foods using multiplex polymerase chain reaction coupled with high-performance liquid chromatography method. *Eur Food Res Technol* 234:649–654. <https://doi.org/10.1007/s00217-012-1671-5>
- Milavec M, Dobnik D, Yang L, Zhang D, Gruden K, Žel J (2014) GMO quantification: valuable experience and insights for the future. *Anal Bioanal Chem* 406:6485–6497. <https://doi.org/10.1007/s00216-014-8077-0>
- Morisset D, Štebih D, Milavec M, Gruden K, Žel J (2013) Quantitative analysis of food and feed samples with droplet digital PCR. *PLoS One* 8:e62583. <https://doi.org/10.1371/journal.pone.0062583>
- Nadal A, Coll A, La Paz JL et al (2006) A new PCR-CGE (size and color) method for simultaneous detection of genetically modified maize events. *Electrophoresis* 27:3879–3888. <https://doi.org/10.1002/elps.200600124>
- Pakpour S (2012) A multi-criteria decision-making approach for comparing sample preservation and DNA extraction methods from swine feces. *Am J Mol Biol* 2:159–169. <https://doi.org/10.4236/ajmb.2012.22018>
- Pennisi E (2017) Pocket-sized sequencers start to pay off big. *Science* (80-). <https://doi.org/10.1126/science.356.6338.572>
- Pla M, Nadal A, Baeten V, Bahrtdt C, Berben G, Bertheau Y, Coll A, van Dijk JP, Dobnik D, Fernandez Pierna JA, Gruden K, Hamels S, Holck A, Holst-Jensen A, Janssen E, Kok EJ, la Paz JL, Laval V, Leimanis S, Malcevski A, Marmioli N, Morisset D, Prins TW, Remacle J, Ujhelyi G, Wulff D (2013) New multiplexing tools for reliable GMO detection. In: Bertheau Y (ed) Genetically modified and non-genetically modified food supply chains: co-existence and traceability. Wiley-Blackwell, Oxford, pp 333–366. <https://doi.org/10.1002/9781118373781.ch19>
- Potapov V, Ong JL (2017) Examining sources of error in PCR by single-molecule sequencing. *PLoS One* 12:e0169774. <https://doi.org/10.1371/journal.pone.0169774>
- Randhawa GJ, Chhabra R, Singh M (2009) Multiplex PCR-based simultaneous amplification of selectable marker and reporter genes for the screening of genetically modified crops. *J Agric Food Chem* 57: 5167–5172. <https://doi.org/10.1021/jf900604h>
- Randhawa GJ, Chhabra R, Singh M (2010) Decaplex and real-time PCR based detection of MON531 and MON15985 Bt cotton events. *J Agric Food Chem* 58:9875–9881. <https://doi.org/10.1021/jf100466n>
- Remund KM, Dixon DA, Wright DL, Holden LR (2001) Statistical considerations in seed purity testing for transgenic traits. *Seed Sci Res* 11:101–119. <https://doi.org/10.1079/SSR200166>
- Takabatake R, Koiwa T, Kasahara M et al (2011) Interlaboratory validation of quantitative duplex real-time PCR method for screening analysis of genetically modified maize. *Food Hyg Saf Sci (Shokuhin Eiseigaku Zasshi)* 52:265–269. <https://doi.org/10.3358/shokueishi.52.265>
- Van Den Bulcke M, De Schrijver A, De Bernardi D et al (2007) Detection of genetically modified plant products by protein strip testing: an evaluation of real-life samples. *Eur Food Res Technol* 225:49–57. <https://doi.org/10.1007/s00217-006-0381-2>
- Van den Bulcke M, Lievens A, Barbau-Piednoir E et al (2010) A theoretical introduction to “Combinatory SYBR@Green qPCR Screening”, a matrix-based approach for the detection of materials derived from genetically modified plants. *Anal Bioanal Chem* 396: 2113–2123. <https://doi.org/10.1007/s00216-009-3286-7>
- Vigani M, Raimondi V, Olper A (2012) International trade and endogenous standards: the case of GMO regulation. *World Trade Rev* 11: 415–437
- Volk H, Piskernik S, Kurinčić M et al (2014) Evaluation of different methods for DNA extraction from milk. *J Food Nutr Res* 53:97–105
- Waiblinger H-U, Ernst B, Anderson A, Pietsch K (2007) Validation and collaborative study of a P35S and T-nos duplex real-time PCR screening method to detect genetically modified organisms in food products. *Eur Food Res Technol* 226:1221–1228. <https://doi.org/10.1007/s00217-007-0748-z>
- Wang C, Li R, Quan S, Shen P, Zhang D, Shi J, Yang L (2015) GMO detection in food and feed through screening by visual loop-mediated isothermal amplification assays. *Anal Bioanal Chem* 407:4829–4834. <https://doi.org/10.1007/s00216-015-8652-z>
- Willems S, Fraiture M-A, Deforce D, de Keersmaecker SCJ, de Loose M, Ruttink T, Herman P, van Nieuwerburgh F, Roosens N (2016) Statistical framework for detection of genetically modified organisms based on next generation sequencing. *Food Chem* 192:788–798. <https://doi.org/10.1016/j.foodchem.2015.07.074>