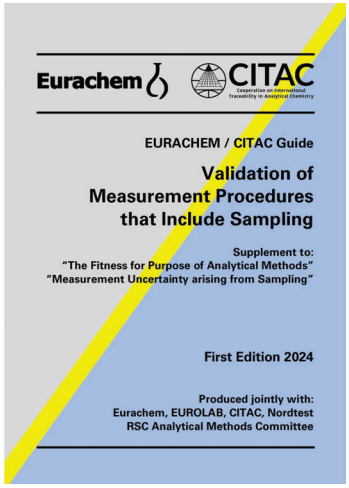
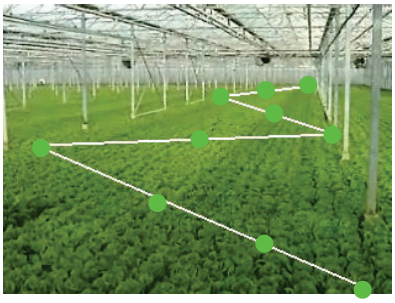
# Validation of Measurement Procedures that Include Sampling (VaMPIS) – New Eurachem Guidance

Prof. Michael H Ramsey
Chair of Eurachem UfS Working Group
School of Life Sciences,
University of Sussex, Brighton, UK
m.h.ramsey@sussex.ac.uk





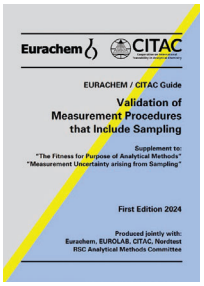


*Eurachem Workshop: Complex Matrices: Applications, Laboratory Standards, and Accreditation, Bucharest, Romania - May 26 – 27, 2025*

University of Sussex

## Overview of Talk

- Rationale for VaMPIS:- Sampling is part of the measurement process
- Validation of measurement procedures that includes sampling (VaMPIS)
  - Not just the analytical component in isolation
    - traditionally assessed using 7 method performance characteristics
  - Judge the fitness for purpose (FFP) of measurement procedures using…
  - Uncertainty (U) of measurement values – *as the key metric*
    - including contributions from sampling (UfS) as well as from analysis (UfA)
  - Estimation of UfS (& MU) - Mainly using Duplicate Method
  - Example for an *ex situ* measurement procedure
  - Management issues in implementation
- Conclusions




University of Sussex

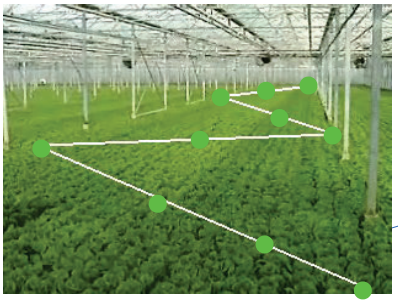
# *Rationale of VaMPIS:*
# Sampling is part of the measurement process

- Sampling really the first step in the measurement process
- *In situ* measurement techniques reveal this
  - Place the sensor→ make measurement = taking a sample
  - Uncertainty from sampling produces MU in measurement
- Physical sample preparation (in field or lab)
  - e.g. filter, acidify, dry, store, sieve, grind, split
  - is also part of the measurement process
  - and potentially important source of MU
  - include in the validation and QC processes (often omitted by labs)


siebtechnik-tema.com

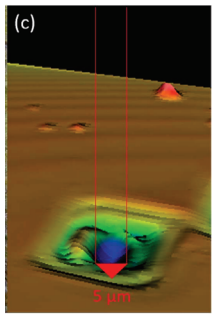
University of Sussex

# The measurement process – including Sampling and Sample Preparation



Sampling Target (at macro scale)
*Portion of material, at a particular time, that a sample is intended to represent*


| Process step | Form of material | Description of process step |
|---|---|---|
| Sampling | Sampling Target | Collection of a single sample, or several increments combined into composite sample |
| | Primary Sample | Comminution and/or splitting |
| | Sub-sample | Further comminution and/or splitting |
| Physical sample preparation | Laboratory sample | Physical preparation, e.g. drying, sieving, milling, splitting, homogenisation |
| | Test sample | Selection of test portion for chemical treatment preceding chemical analysis |
| | Test portion | Chemical treatment leading to analytical determination |
| Analysis | Test solution | Analytical determination of analyte concentration |

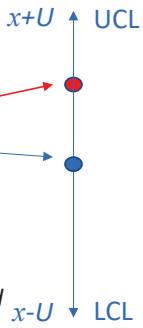



Sampling Target at micron scale

**Hence need to validate the whole measurement procedure = VaMPIS**

University of Sussex

# Measurement Uncertainty (MU) – *key metric to judge FFP*




- Historially: MU (*U*) is *'an estimate attached to a test results (x)….*
  *which characterises the range of values within which the true value is asserted to lie'* [1]
  - 'True value' equivalent to '**Value of the Measurand**' in more recent definitions
  - *Parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand.* [2]
  - UCL = Upper Confidence Limit, LCL = Lower Confidence Limit.
  - **Confidence Interval** (CI) is between LCL and UCL
- Includes both Random effects (e.g. precision) and Systematic effects (e.g. bias)
- MU arises from <u>all</u> steps in measurement (e.g. sampling & physical sample prep.) - *in ISO/IEC 17025*
  - Doesn't <u>assume</u> that samples are 'correct' – hence 'representative'
    - traditional approach to Sampling Quality
- MU is key parameter of measurement (and sampling) quality – reflects contribution from all steps

[1] Historic definition of MU from ISO 3534-1: 1993 Statistics – Vocabulary and Symbols, International Organization for Standardization, Geneva
[2] JCGM 100 (2008) / ISO/IEC Guide 98-3:2008

University of Sussex

# Statistical model
## for ***Empirical*** estimation of uncertainty - One Sampling Target

$$x = X_{true} + \varepsilon_{sampling} + \varepsilon_{analytical}$$

$x$ = ***measured*** **value** of the analyte concentration in one sampling target

$X_{true}$ = ***true*** **value** of the analyte concentration in the sampling target

$\varepsilon_{sampling} + \varepsilon_{analytical}$ = effects on measured concentration from sampling and analysis

**Variance** (standard deviation squared) **of measurement value** = $\sigma^2_{meas}$

$$\sigma^2_{meas} = \sigma^2_{sampling} + \sigma^2_{analytical}$$

$\sigma^2_{sampling}$ is the between-sample variance on one target (largely due to analyte heterogeneity)
$\sigma^2_{analytical}$ is the between-analysis variance on one sample (as Repeatability)

For ***estimates*** of variance, we have:

$$s^2_{meas} = s^2_{sampling} + s^2_{analytical}$$

University of Sussex

# Statistical model

## for ***Empirical*** estimation of uncertainty - Multiple Sampling Targets

**Multiple sampling targets ( n > 8) are needed for more realistic estimate of MU & UfS – using SPT**

$$x = X_{true} + \varepsilon_{target} + \varepsilon_{sampling} + \varepsilon_{analytical}$$

$\varepsilon_{target}$ represents the variation of concentration between the targets and has variance $\sigma^2_{between\ -target}$ .

**Variance of measurement value** = $\sigma^2_{meas} = \sigma^2_{sampling} + \sigma^2_{analytical}$

$$\sigma^2_{total} = \sigma^2_{between\ -target} + \sigma^2_{sampling} + \sigma^2_{analytical}$$

For our estimates of variances, we have:-

$$s^2_{total} = s^2_{between-target} + s^2_{sampling} + s^2_{analytical}$$

$$s^2_{meas}$$

University of Sussex

# How MU is expressed & reported

- MU usually expressed using standard deviation (*s*), e.g.:-

1. Standard uncertainty (*u*)

   $u = s_{meas}$ (often = $s_{analytical}$)

2. Expanded uncertainty (*U*)

   $U = ks_{meas} = 2s_{meas}$

   with coverage factor (*k*) of 2 for 95% confidence

   - may need k > 2 for *U* based upon small number of samples*

3. Expanded relative uncertainty (*U'*), relative to measurement value ($x$)

$$U' = 100\frac{2s_{meas}}{x}\%$$

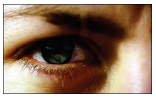
MU can also be expressed as a Confidence Interval, e.g. LCL – UCL

- or as an Uncertainty Factor (if *U'* large (>20%) or log-normally distributed)

*Rostron P.D., Fearn T., and Ramsey M.H. Improved coverage factors for Expanded Measurement Uncertainty calculated from two estimated variance components. Accred Qual Assur (2024) 29:225–230. https://doi.org/10.1007/s00769-024-01579-w - used in RANOVA4 software

University of Sussex

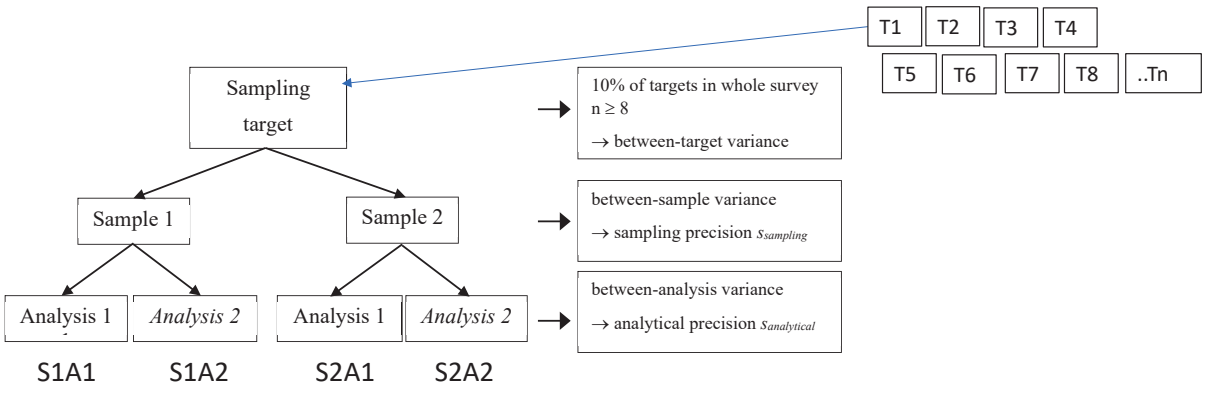
# Duplicate Method of UfS Estimation – General Principles




- Duplication is most cost-effective form of replication
  - Apply to both duplicate samples and duplicate chemical analyses
    - using two-stage nested experimental design (balanced or unbalanced)
  - But can have large confidence interval of resulting estimates of MU
  - Unless it is applied to at least 8 sampling targets (ideally more, e.g. 20)
- Realistic taking of duplicate samples is crucial
  - Not just the splitting of a single sample
- Take duplicate samples independently by fresh interpretation of the sampling procedure
  - How far away (in space or time) might duplicate sample be taken? Reflects..
    - ambiguity in sampling procedure
    - spatial uncertainty in the surveying device in use
    - Example below for *ex situ* measurement of Nitrate in lettuce (UfS-A1, VaMPIS-B1)

University of Sussex

# Estimation of MU (including UfS) Using Duplicate Method – Full Balanced Design

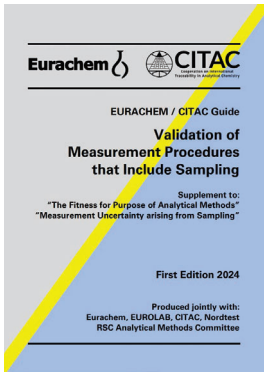



- Usually uses this full balanced experimental design *(unbalanced - no S2A2 - reduces cost)*
- 8 typical Sampling Targets chosen
- Only requires one 'sampler' (or measurement scientist)
  - Can be improved using multiple 'samplers' - using SPT results *(see VaMPIS & UfS Guide)*
- Explain Duplicate Method for Case Studies – followed by ANOVA
  - Applicable to Validate both *ex situ* and *in situ* measurement methods – *flow chart*

University of Sussex

## Validation of Measurement Procedures Including Sampling (VaMPIS)

*- use the uncertainty of the final measurement value to judge its FFP*
*- and hence the FFP of the whole measurement procedure*





VaMPIS applied either:-

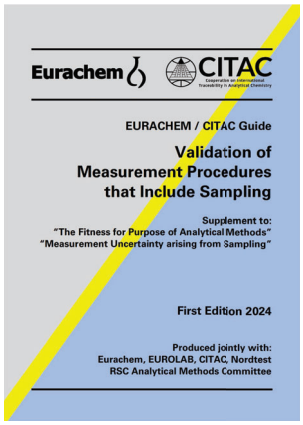
- *sequentially* (with previously validated analytical procedure)
- or *simultaneously* (sampling and analysis together)
- on *ex situ* measurement procedures
- or *in situ* measurement procedures

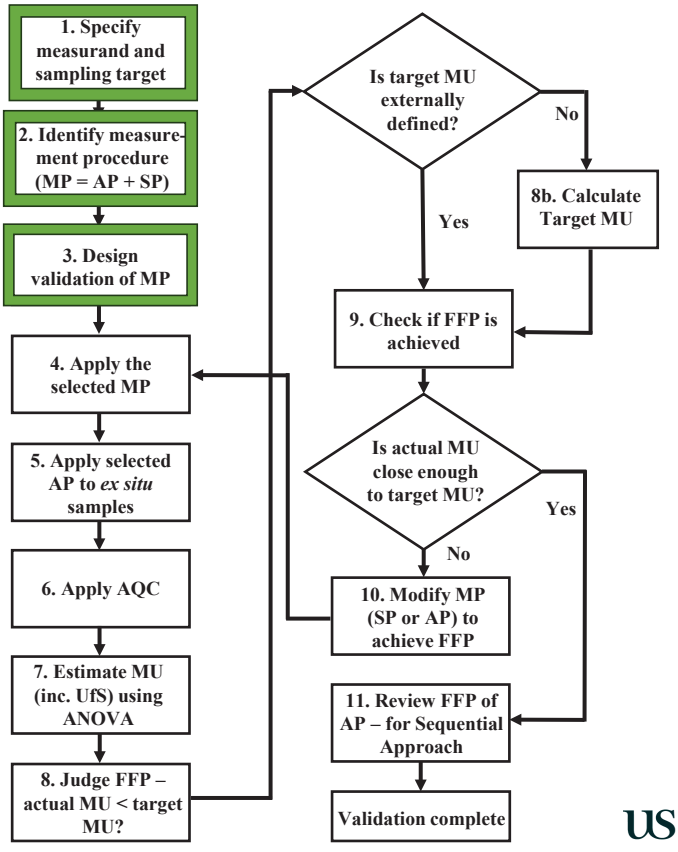
University of Sussex

## Validation of Measurement Procedures Including Sampling (VaMPIS)

*- Flow Chart*

*- 11 steps*
*- best explained with an example*

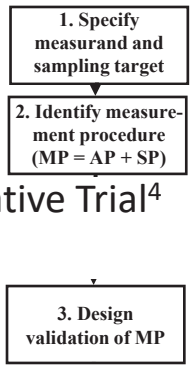



University of Sussex

# VaMPIS applied to Nitrate Concentration in Lettuce
## - applied sequentially, to *ex situ* measurement procedure

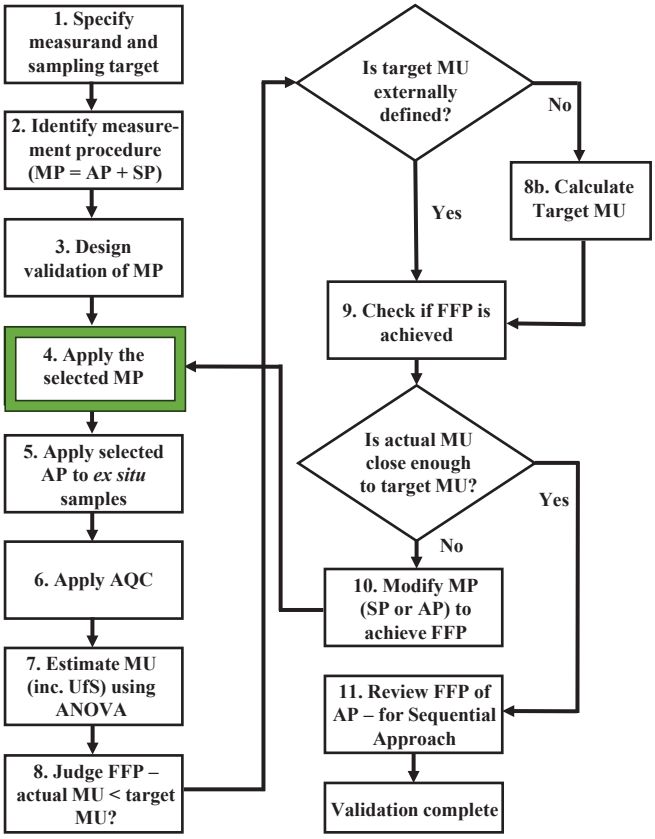
- EU threshold 4500 mg $kg^{-1}$ for nitrate concentration of Sampling Target[1]
  - i.e. ~ 12,000 – 20,000 heads in each bay/batch/target
- Current EU Sampling Procedure[2] specifies taking 10 heads (increments)
  - to make a single **composite sample** from each Sampling Target
- Analytical Procedure/method (HPLC[3]) **already validated** using Collaborative Trial[4]
  - $U_{analysis}$ around 6% at that validation ($RSD_{Reproducibility}$ = ~ 3%)
- Need to validate the whole measurement procedure
  - including sampling & sample preparation
- MU is key metric that affects compliance decisions
  - MU is affected by (and reflects) all 7 performance characteristics for measurement procedure
- Judge FFP of measurement procedure by the Actual MU
  - - is it close enough to Target MU?




1. Commission Regulation (EC) No 563/2002 of 2 April 2002 amending Regulation (EC) No 466/2001
2. European Directive 79/700/EEC. OJ L 207, 15.8.1979, p26.
3. BS EN 12014-2:1997, Foodstuffs. Determination of nitrate and/or nitrite content. General considerations
4. Farrington et al.,(2006), Journal of the Association of Public Analysts (Online), 34, 1-11

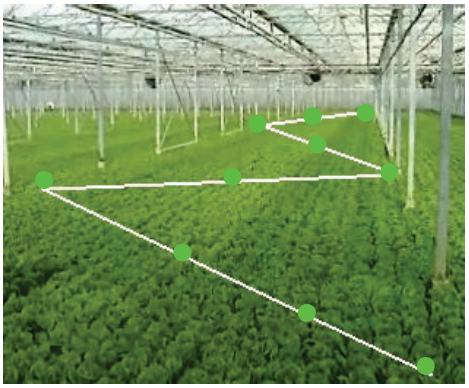
University of Sussex

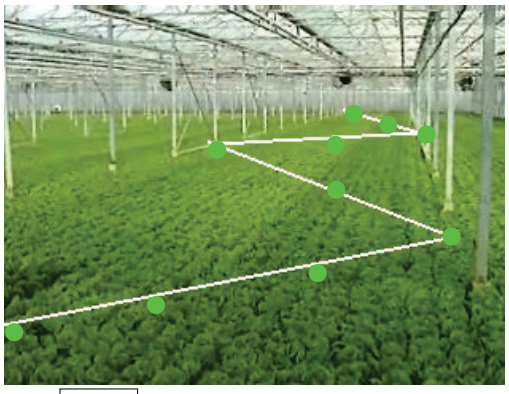
**Validation of Measurement Procedures Including Sampling (VaMPIS)**
*- Flow Chart*

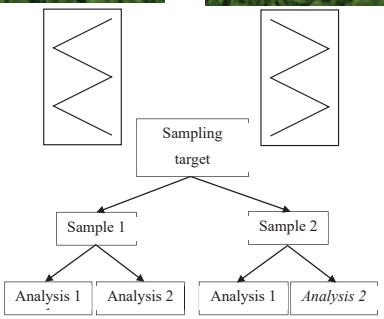



University of Sussex

# **UfS estimation for Lettuce** using Duplicated ‘W’ Sampling Design



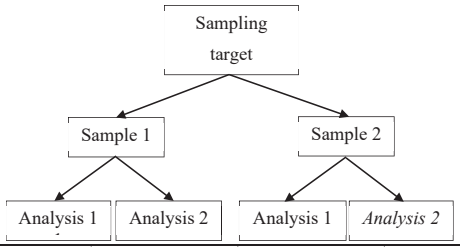





Duplicate sample is equally likely interpretation of ‘W’ design

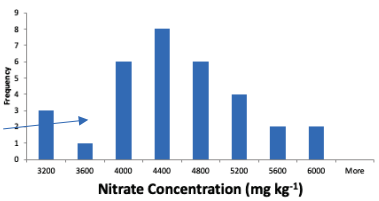
University of Sussex

# Estimating UfS (and MU) for Nitrate in Lettuce




| Sample target | S1A1 | S1A2 | S2A1 | S2A2 |
|---|---|---|---|---|
| | Nitrate concentration (mg kg$^{-1}$) | | | |
| A | 3898 | 4139 | 4466 | 4693 |
| B | 3910 | 3993 | 4201 | 4126 |
| C | 5708 | 5903 | 4061 | 3782 |
| D | 5028 | 4754 | 5450 | 5416 |
| E | 4640 | 4401 | 4248 | 4191 |
| F | 5182 | 5023 | 4662 | 4839 |
| G | 3028 | 3224 | 3023 | 2901 |
| H | 3966 | 4283 | 4131 | 3788 |

Distribution approx. Normal + <10% outliers Needs Robust ANOVA

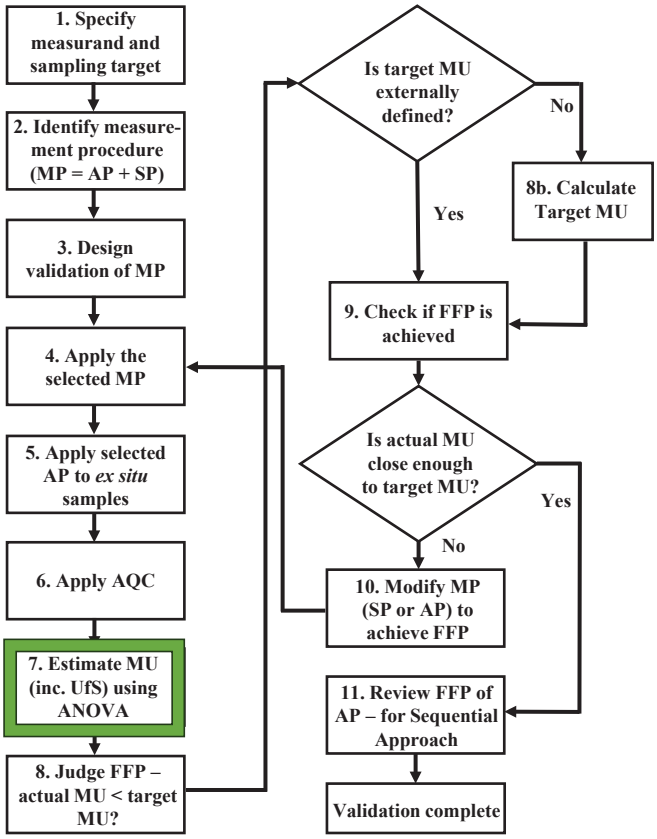



Lyn, J.A., Palestra, I.M., Ramsey, M.H., Damant, A.P. and Wood, R. (2007) Modifying uncertainty from sampling to achieve fitness for purpose: a case study on nitrate in lettuce *Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement,* 12, 67-74

University of Sussex

## Validation of Measurement Procedures Including Sampling (VaMPIS)

- *Flow Chart*




University of Sussex

# RANOVA3 output for Uncertainty estimation

*Input the 32 measurement values from balanced design*

**Robust ANOVA**

| | | | | |
|---|---|---|---|---|
| Mean | 4408.3 | | | |
| Total Sdev | 670.58 | | | |
| | Btn Target | Sampling | Analysis | Measure |
| Standard dev | 565.4 | 319.05 | 167.94 | 360.55 |
| % of total vari | 71.09 | 22.64 | 6.27 | 28.91 |
| Expanded relative uncertair | 14.47 | 7.62 | | 16.36 |

$U'_{anal}$ = 7.6% *(as repeatability)*

*– but ignoring sampling*

*Underestimates MU of 16%*

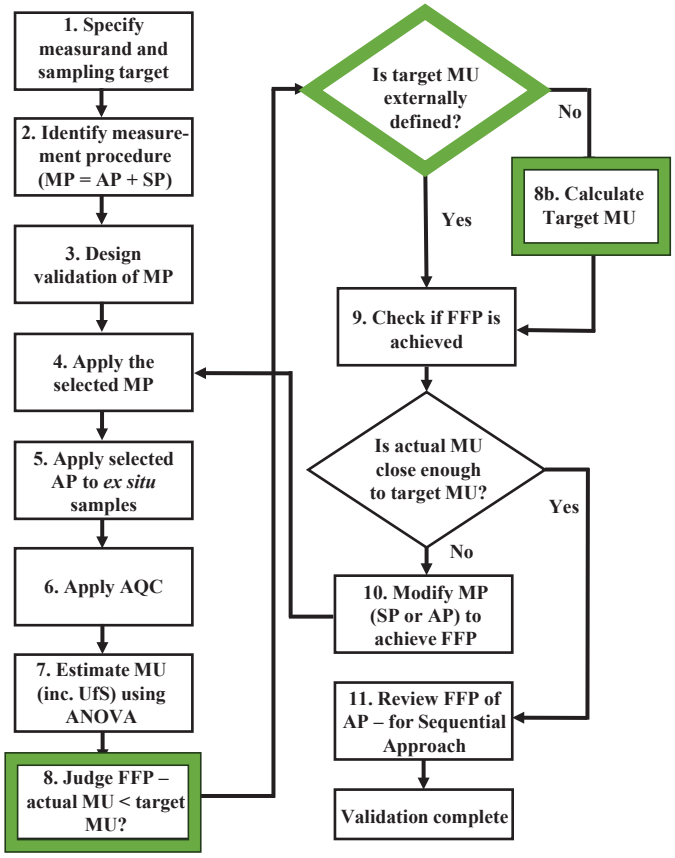
$u$ = 361 mg kg$^{-1}$

$U'$ in % (k=2) = 16%

*RANOVA3 available free from AMC Software website*

University of Sussex

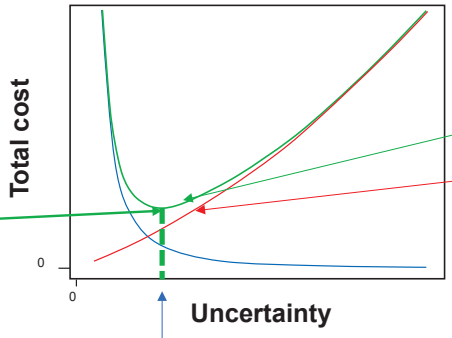
**Validation of Measurement Procedures Including Sampling (VaMPIS)**
*- Flow Chart*




University of Sussex

# Judge FFP by comparing Actual MU against Target MU

- Target MU (inc. UfS + UfA) - can be **Option (1) set externally** (e.g. arbitrary 20%, 16% < 20% so FFP), **or**
- **Option (2): Calculate,** for example at….
- **Optimal MU*** that minimises the overall cost (including the consequences of incorrect decisions)




**Minimum cost where measurements are FFP**
- **at Optimal MU**
- **set as Target MU**

Sum of both costs

Cost of incorrect decisions
e.g. Lettuce €5280 (12,000 heads at €0.44)
- for false positive decision (unnecessary disposal)

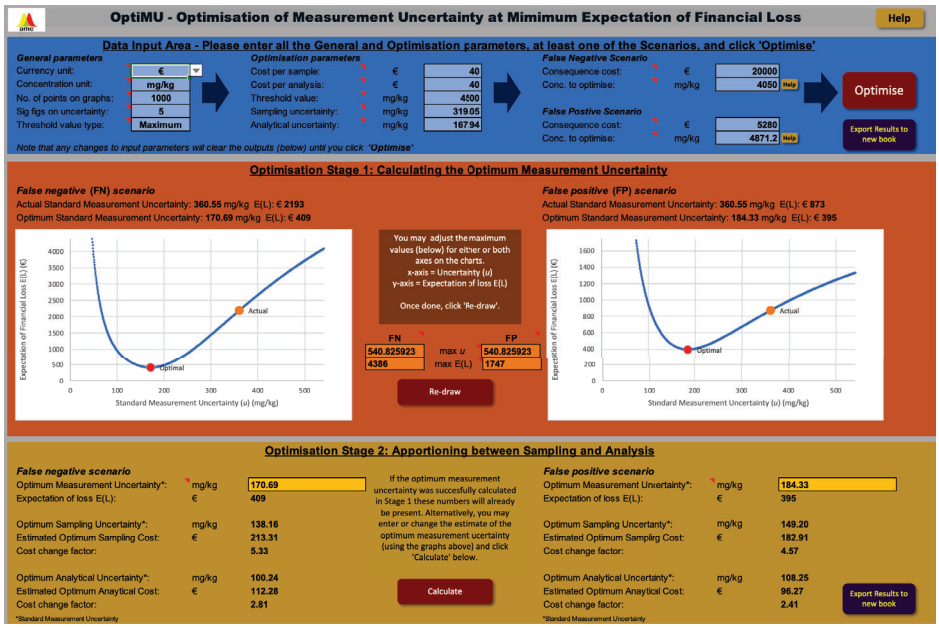
Cost of measurement
e.g. Lettuce €40 per sample, €40 per analysis

- Knowing UfS & UfA, can judge how Target MU (however set) can be achieved most cost-effectively by:
  - *Either* Spending more (or less) on **chemical analysis** (e.g. more precise technique)
  - *Or* Spending more (or less) on **sampling** (e.g. taking more increments)

- Theory explained in Appendix B of VaMPIS, two worked examples in Appendix A
- Apply using OptiMU* software (will be available free from AMC website)

University of Sussex

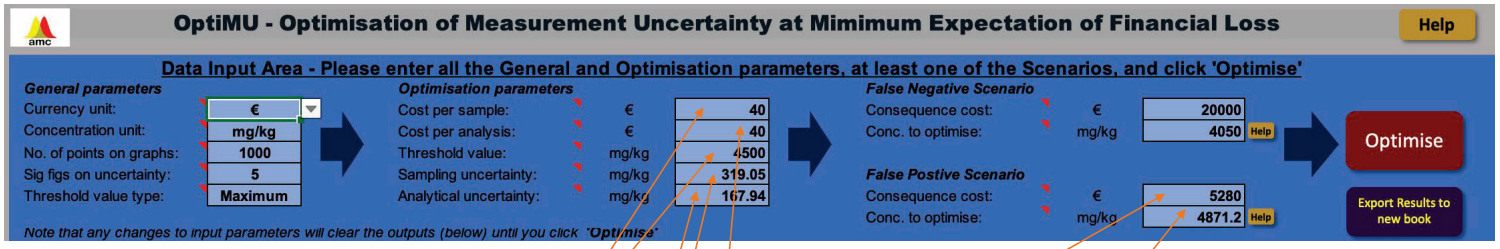
# Output of OptiMU – for Calculating Target MU




- Consider each of the three sections in turn:-
- Applied to this case study

University of Sussex

# Calculating Target MU using OptiMU – *Input data*



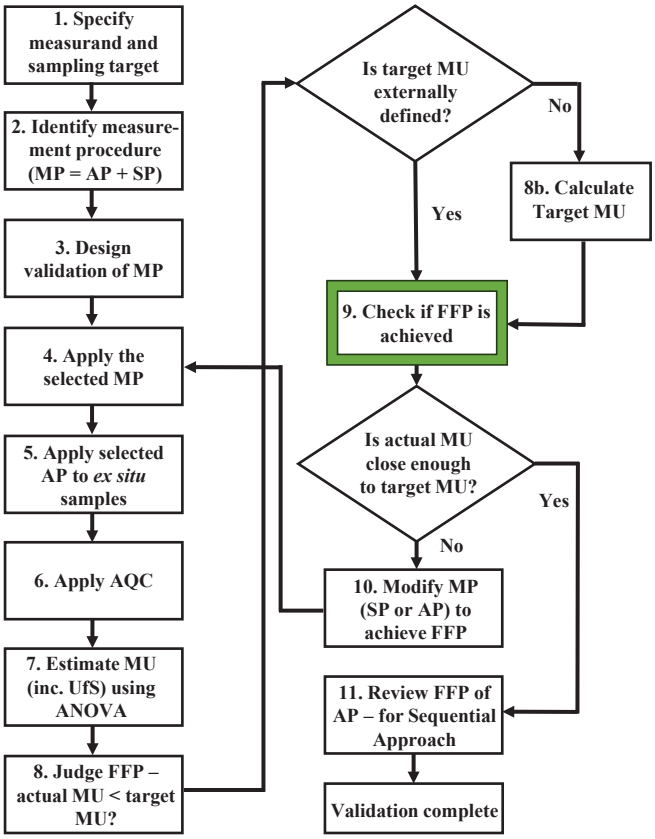

- Threshold value e.g. 4500 mg $kg^{-1}$
- MU values from ANOVA output
- Cost values: sampling and analysis as charged by the samplers/lab
- Consequence costs, e.g. value of 20,000 lettuce heads for false positive
- Concentration at which to optimize, e.g. 4871.2 mg $kg^{-1}$

  = Minimum concentration which would indicate that nitrate concentration was greater than threshold

University of Sussex

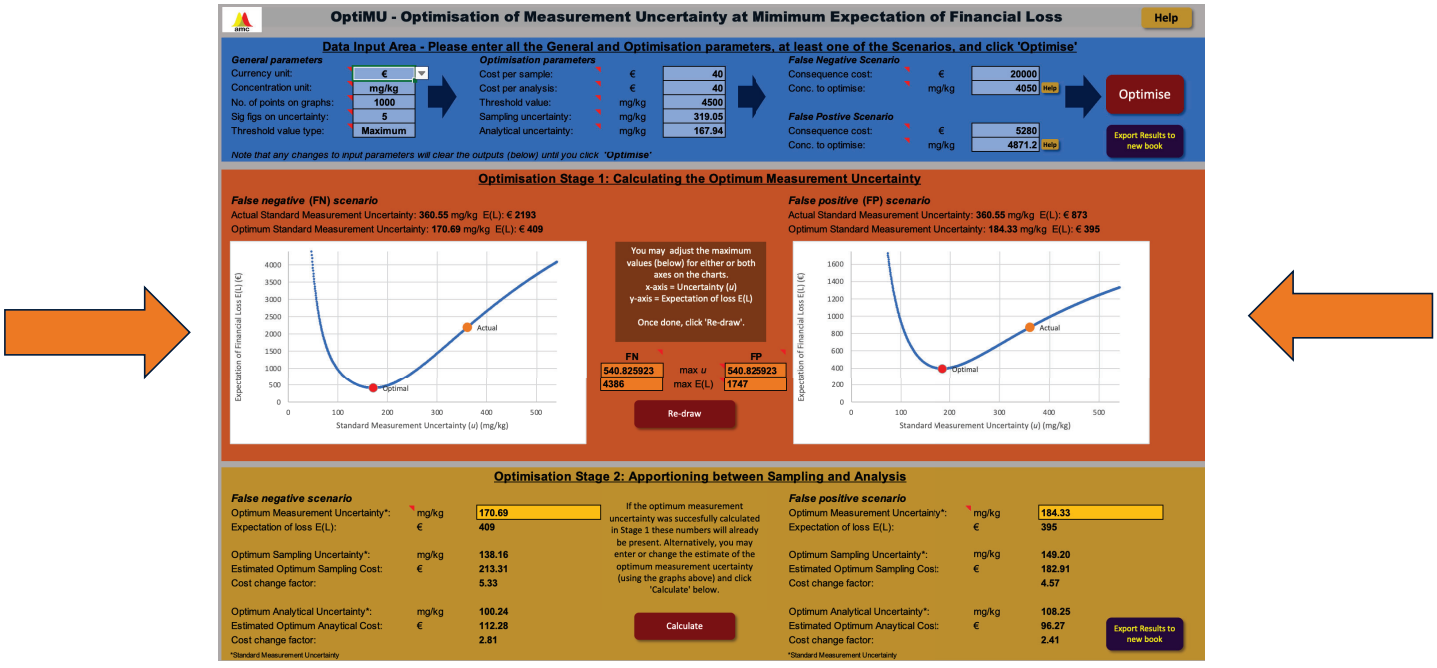
## Validation of Measurement Procedures Including Sampling (VaMPIS)

*- Flow Chart*



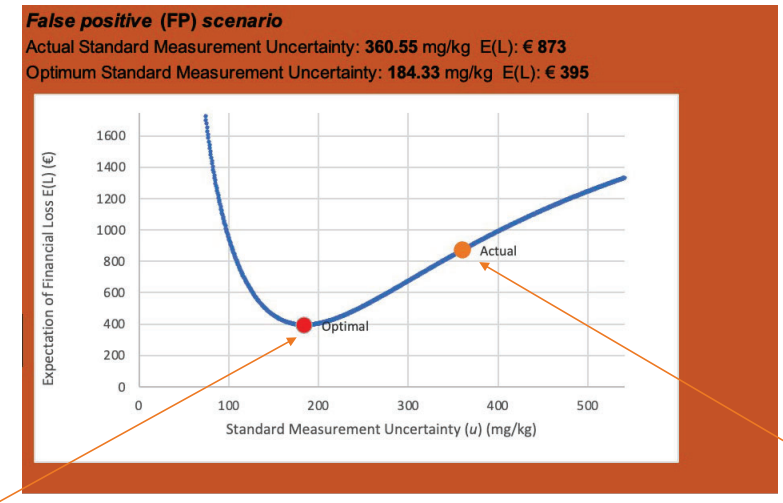

University of Sussex

# Output of OptiMU – for Setting Target MU




University of Sussex

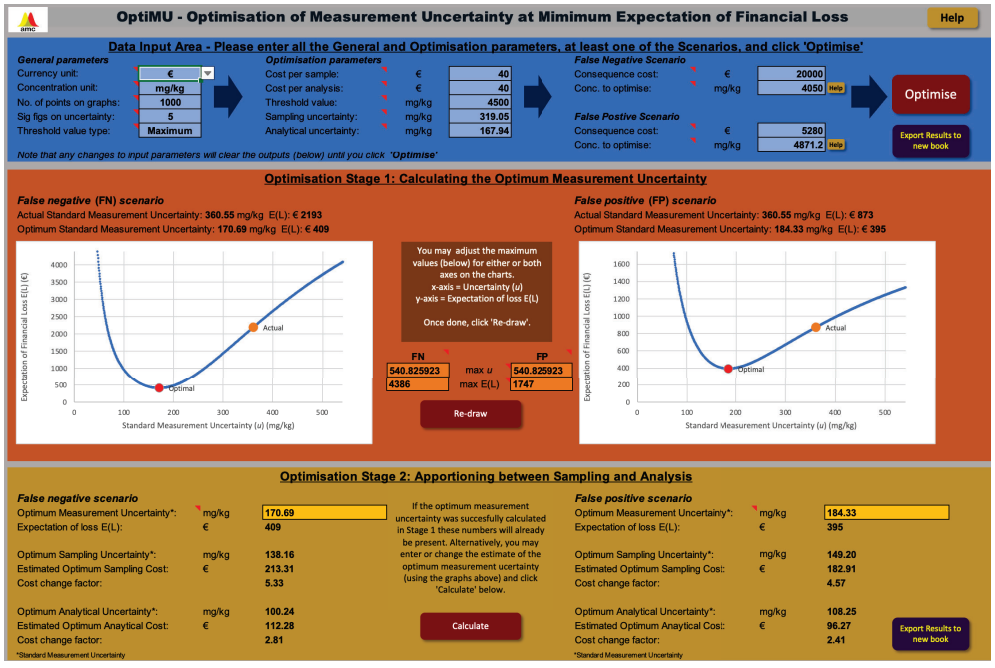
# Setting Target MU - OptiMU- Optimisation (Part 1)




- Optimal MU calculated at minimum of loss function
- *= 184 mg kg$^{-1}$ For **False Positive** – under half of the Actual MU (= 360 mg/kg$^{-1}$)*
- *Loss at Optimal MU = €395 – under half of the €873 at Actual MU*
- ***Measurement Procedure is therefore <u>NOT Fit for Purpose</u>***

University of Sussex

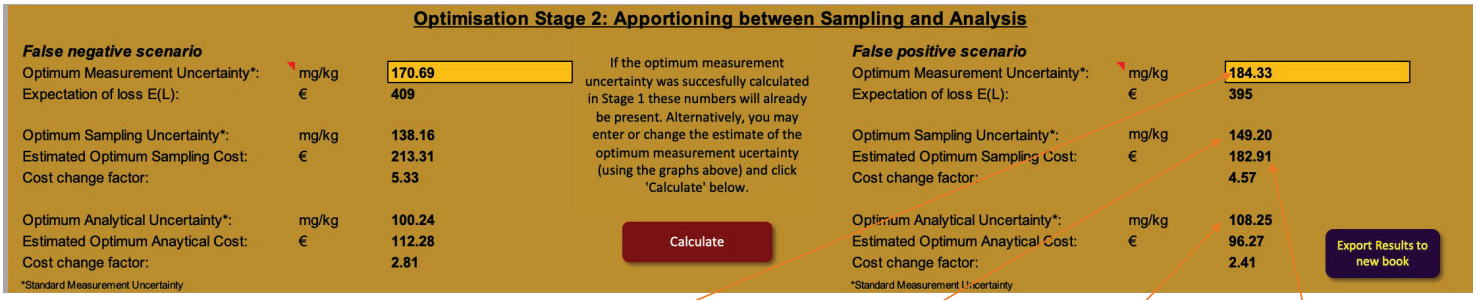
# Output of OptiMU – for Setting Target MU




University of Sussex

# Apportioning MU between Sampling and Analysis

## OptiMU-(Part 2)



**Optimisation Stage 2: Apportioning between Sampling and Analysis**

| *False negative scenario* | | | | *False positive scenario* | | |
|---|---|---|---|---|---|---|
| Optimum Measurement Uncertainty*: | mg/kg | 170.69 | If the optimum measurement uncertainty was succesfully calculated in Stage 1 these numbers will already be present. Alternatively, you may enter or change the estimate of the optimum measurement ucertainty (using the graphs above) and click 'Calculate' below. | Optimum Measurement Uncertainty*: | mg/kg | 184.33 |
| Expectation of loss E(L): | € | 409 | | Expectation of loss E(L): | € | 395 |
| Optimum Sampling Uncertainty*: | mg/kg | 138.16 | | Optimum Sampling Uncertainty*: | mg/kg | 149.20 |
| Estimated Optimum Sampling Cost: | € | 213.31 | | Estimated Optimum Sampling Cost: | € | 182.91 |
| Cost change factor: | | 5.33 | | Cost change factor: | | 4.57 |
| Optimum Analytical Uncertainty*: | mg/kg | 100.24 | Calculate | Optimum Analytical Uncertainty*: | mg/kg | 108.25 |
| Estimated Optimum Anaytical Cost: | € | 112.28 | | Estimated Optimum Anaytical Cost: | € | 96.27 |
| Cost change factor: | | 2.81 | | Cost change factor: | | 2.41 |

*Standard Measurement Uncertainty

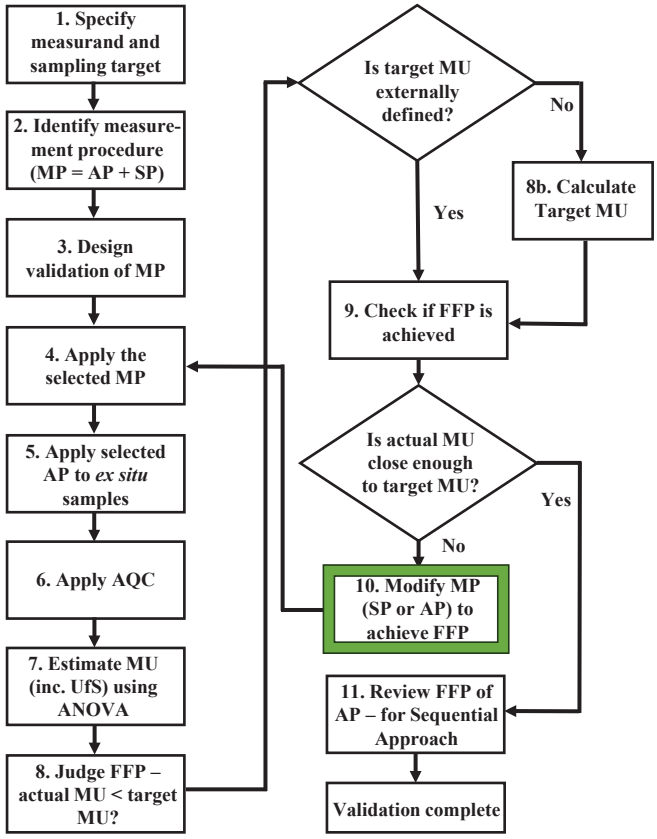
*Standard Measurement Uncertainty

Export Results to new book

- Part 2 sub-divides Optimal MU (184 $mg\ kg^{-1}$) into UfS and UfA – *to gives optimal values for both*
- UfS needs to be reduced from 319 to 149 mg $kg^{-1}$ to achieve Target (FFP)
- Can be achieved by increasing cost of sampling (x 4.6) from €40 to €183
  - e.g. by taking 40-head (rather than 10-head) composite sample
  - *Model from Sampling Theory predicts 2-fold drop in UfS for 4-fold increase sample mass*
- UfA could also be reduced (by 36%) from 168 to 108 mg $kg^{-1}$ to help achieve Target (FFP)
- To achieve this cost of analysis could be increased (x 2.4) from €40 to €96 e.g. more precise method
- As UfS accounts for 78% of MU, Target MU best approached by reducing UfS x2

University of Sussex

**Validation of Measurement Procedures Including Sampling (VaMPIS)**
*- Flow Chart*

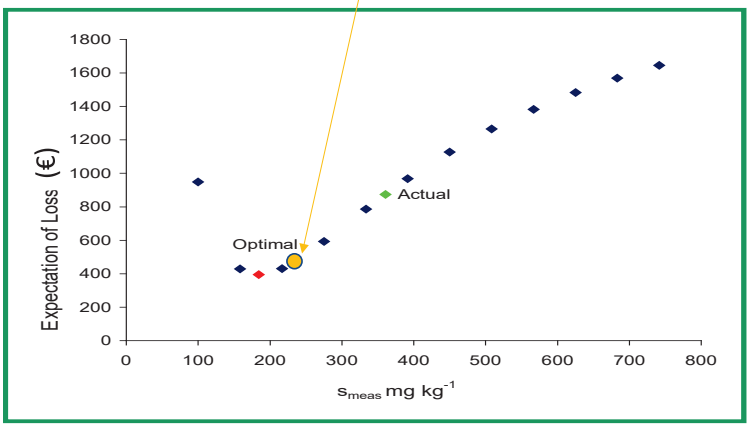



University of Sussex

# Reducing the Uncertainty – to achieve FFP

- **Increasing number of increments from 10 to 40 heads**
- Reduced $s_{samp}$ from 319 to 177 mg $kg^{-1}$ - by a factor of x 1.8 ( similar to model prediction of x2)
- Reduced MU ($s_{meas}$) from 360 to 244 mg $kg^{-1.}$ (*U'* from 16.4 % to 11.1%)
- Close to the optimal value (184 mg $kg^{-1}$) at similar Cost (~€500 per target, down from €800)
- A**chieves Fitness-for-Purpose (FFP)** = MU that effectively minimises to overall financial loss

Cost ↑



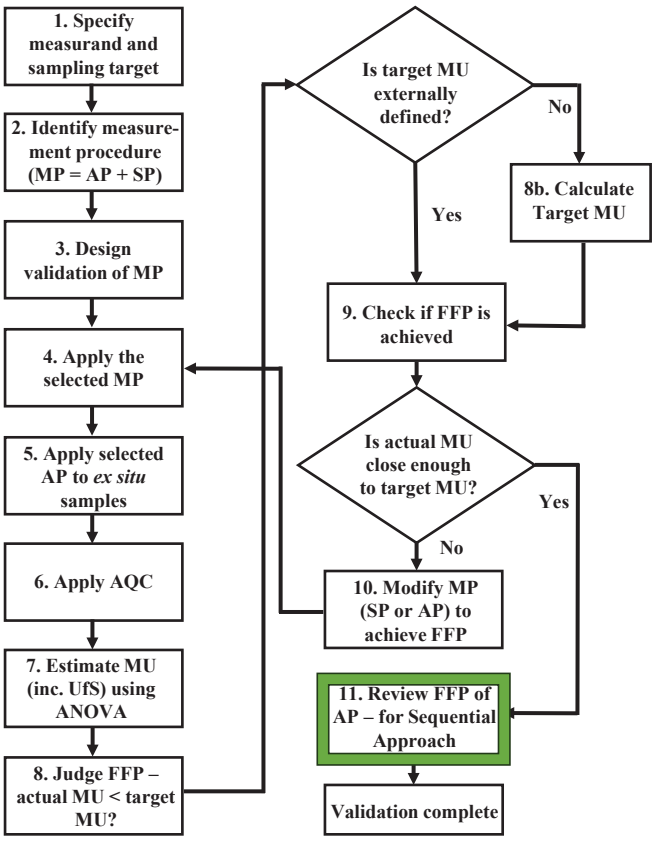

Uncertainty→

Lyn, J.A., Palestra, I.M., Ramsey, M.H., Damant, A.P. and Wood, R. (2007) Modifying uncertainty from sampling to achieve fitness for purpose: a case study on nitrate in lettuce *Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement,* 12, 67-74

University of Sussex

## Validation of Measurement Procedures Including Sampling (VaMPIS)

*- Flow Chart*




University of Sussex

## Review FFP of Analytical Procedure For sequential approach to VaMPIS

- Analytical MU estimated previously (by CT) in isolation = 6.0% - judged as FFP
- Analytical MU estimated from Duplicate Method* = 7.6% *(nominally larger, despite….)*

  ** Excluded, between-day and between operator variability and between-lab bias, but…*

  ** Included heterogeneity of routine test materials (rather than use very homogeneous CRMs or IQC materials to estimate)*

  – Probably not statistically significantly different (e.g. 7.6% has 95% CI from 6.3% to 9.4%, RANOVA4)
- So analytical method/procedure is still FFP by that criterion.

**Using the OU method indicates that:-**

- 10-head sampling procedure not FFP - and that MU should be reduced by:-

(a) Improving the sampling (e.g. taking 40 head-composite samples), or also

(b) Improving analysis (reducing $U_{anal}$ by 36%, from 7.6% to 4.9%)

MU sufficiently reduced with adjusting $U_{samp}$, without need to also reduce $U_{anal}$

University of Sussex

## Management issues in VaMPIS implementation

- Traditionally sampling often undertaken by separate organisation
- Makes implementation of VaMAPIS a challenge
- Need better coordination & communication between lab and samplers
  - Before, during and after VaMPIS:-
- Need for ongoing QC of the whole measurement process
  - Integrated Measurement Quality Control (IMQC)
  - To include QC of both sampling and chemical analysis
- Discussed in Sections 3 & 4 of VaMPIS Guide

University of Sussex

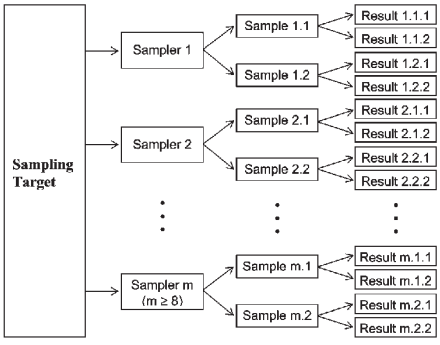
# Conclusions

- New Eurachem Guide: Validation of Measurement Procedures that Include Sampling (VaMPIS)
- Is needed because Sampling is part of the measurement process
- MU is used to judge the FFP of the whole measurement process
  - Summarises the effects on quality of all of the other 7 performance characteristics
  - Is the one characteristic used in compliance decision
- UfS (and hence MU) can be estimated with Duplicate Method *(most practical)*
  - Applicable to any sampling medium: soil, sediment, herbage, waters, gases etc.
  - **Also applicable to *in situ* measurements** (such as PXRF – Example A2 in VaMPIS)
  - Sampling PT (or CT) results can be used to also include between-sampler bias within MU
- VaMPIS can show that (sometimes) to achieve FFP (and hence validation)….
  - sampling (not analysis) needs improvement (e.g. 40-head composite lettuce sample)
- *Questions?*


University of Sussex

# Estimate of Uncertainty using SPT - including Between-Sampler Bias

## - Example using Sampling PT for moisture in butter*




* Ramsey M.H. Geelhoed B, Damant, A.P., Wood, R. (2011) Improved evaluation of measurement uncertainty from sampling by inclusion of between-sampler bias using sampling proficiency testing. Analyst, 136 (7), 1313 – 1321. DOI:10.1039/C0AN00705F.

ANOVA: $U'$ as % on concentration of moisture in butter (200 tons)

≈ Duplicate Method (single sampler) gives $\boldsymbol{U' = 0.39\ \%}$

**SPT (multiple samplers, n=9)** **gives $U' = 0.87\%$**

**- *$U'$ larger* x 2.2* - includes bias between-samplers**

Remove two samplers with potentially non-proficient z-scores (RSz > 3)

**SPT (n=7)** gives $\boldsymbol{U' = 0.69\%}$

- **$U'$ still larger x 1.8**
- **a more reliable estimate of Uncertainty**
- Ideally apply over multiple rounds of SPT, if targets comparable
- e.g. 16 rounds, stack-gas measurement SPT [Coleman et al ,2013, Accred Qual Assur 18:517–524]
- Multiple samplers using one procedure (CTS) better for VaMPIS
- More expensive than Duplicate Method, but sometimes justified

University of Sussex