Using a risk-based approach to Measurement Uncertainty in the clinical flow lab

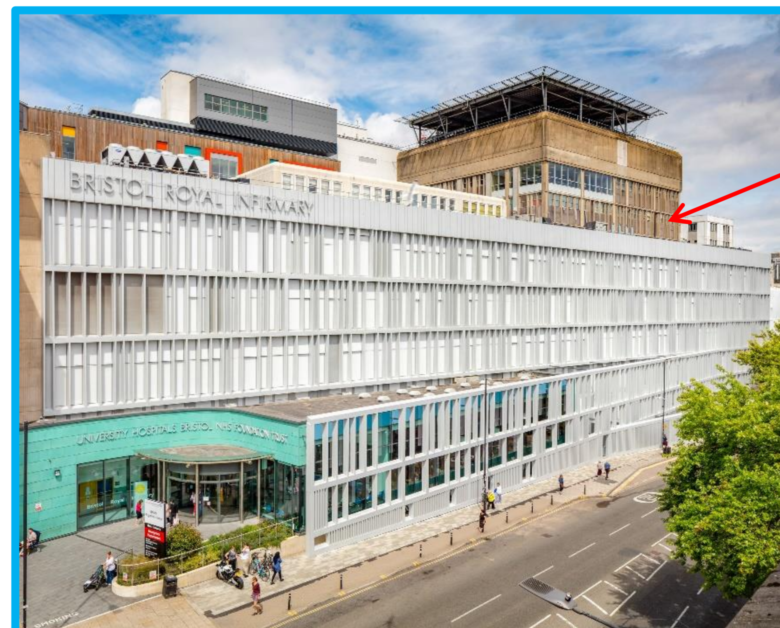
Ulrika Johansson
UHBW SI-HMDS
Bristol

NHS
University Hospitals Bristol and Weston
NHS Foundation Trust

ESCCA
European Society for Clinical Cell Analysis

My background, Disclosures

- Consultant Clinical Scientist at UHBW; a tertiary haemato-oncology centre
- Our Laboratories are ISO15189 accredited since 2017
- Operational lead and flow lab lead for the blood cancer diagnostics unit



Disclosures

BD, support for meetings and workshop

BC, support for meetings and workshop

Milteny Biosciences, support for meetings

Agilent, support for meetings and workshops

Measurement Uncertainty

What does it mean?

Bell, 2001; ISO, 2019; Lee et al., 2015:

MU is a value that indicates the range (or interval) within which the true value should be found, with a stated level of confidence, and it can be presented as an absolute value (with the same unit as the assay result) or a relative one (%)

- **The concept "Measurement Uncertainty" is that of a calculated, numerical value.**
- **It may (in part) be based on precision but goes beyond that (a little more on that later)**

Measurement Uncertainty

A quite new concept and requirement for Flow Cytometry Laboratories

- The application of MU is quite new in clinical laboratories overall, not least in clinical flow cytometry laboratories.

- UK experience:
  - Strict assessors.
  - There should be a MU – can we show our work on MU to them please..
  - if not – what are your plans for this,,,and so forth.

- I asked around – what are others thinking /doing around this?
  - Reached out to UK and ESCCA colleagues

ESCCA Working group on MU

Aim: Produce guidance that can help each individual laboratory – support clinical flow labs

Varied experiences,

Varied approached.

Some countries not yet ISO15189 accredited

Others have been for several years

Some labs accredited to several standards (CAP, ISO)

→ If we are accredited to the same standards (15189) – then it ought to be possible to use a harmonised, acceptable approach.

→ Does interpretation of the ISO standards vary between accreditation bodies?

Catarina Martins
Dan Payne
Dave Bloxham
Eleni Liskens
Evi Yiannaki
Francesco Buccisano
Juan Flores-Montero
Katherina Psarra
Kathryn Clarke
Katy Jones
Maria João Acosta
Neil McIver-Brown
Richard Veyrat-Masson
Thomas Keller
Ulrich Sach
Ulrika Johansson
Wolfgang Kern

What does the Standards say

ISO15189:2022

Evaluation of measurement uncertainty (MU)

7.3.4 Evaluation of measurement uncertainty (MU)

a) The MU of measured **quantity values** shall be evaluated and maintained for its intended use, **where relevant**. The MU shall be compared against performance specifications and documented.

NOTE **ISO/TS 20914 provides details on these activities together with examples**.

b) MU evaluations shall be **regularly reviewed.**

c) For examination procedures where estimation of MU is not possible or applicable, **the rationale for exclusion from MU estimation shall be documented**.

ISO15189:2022

ESCCA Working group: Key question to answer

Aim: Produce guidance that can help each individual laboratory – support clinical flow labs

When *must* we have a numerical, calculated MU?

ISO speaks of “Quantitative Values”

MRD assays?

- Provide numerical result (% disease)
- Provide therapy guiding results

PNH assay?

- Provide numerical result (% disease)
- Gold standard status of flow – what other tests supports the flow findings?

Absolute enumeration assays?

ESCCA Working group: Key question to answer

Aim: Produce guidance that can help each individual laboratory – support clinical flow labs

Is a calculated numerical MU value of benefit to the patient / to the quality of the service?

Does it increase quality – does it reduce the risk of wrong results?

For example – CD34 assay: If MU was 2 cells/ul,

A result of 8 could range from 6 to 10 – and at 10, a decision of proceeding with harvest could potentially be taken…

Whereas, for CLL-like MBL versus CLL, therapy is guided not by that cut-off, but other factors.

If you decide to estimate MU: How would you calculate that value (precision,,,,) – do you have relevant data to do so?

How would you monitor the MU for , for example, a B-ALL MRD assay

- (all have slightly different LAIP)
- Precision of a given patient's MRD at 0.05% may be different from that of another patient's disease at , say 0.1%

ESCCA
European Society for Clinical Cell Analysis

ESCCA Working group on MU

Aim: Produce guidance that can help each individual laboratory – support clinical flow labs

We looked to the H62 and other high-level guidelines, and considered

Flow assays may (mostly) be Qualitative or Semi-quantitative.

Absolute enumeration assays are considered semi-quantitative

→ Where absolute enumeration results are provided, and if EQA and (commercial) IQC are available: a numerical MU (calculated) is feasible.

- → And reasonable, and relevant to do.
- → Other assays are considered Qualitative.

→ Therefore, our consensus for when a numerical MU is recommended is: Only for absolute enumeration assays, where IQC (to enable repeated measurements of the exact same sample, in our own laboratory) and EQA samples (to enable repeat measurements on the same sample, by several different laboratories) are available.

(manuscript submitted)

So how does one calculate MU?

ISO advices the GUM " top down approach"

TECHNICAL SPECIFICATION

ISO/TS 20914

First edition
2019-07

Medical laboratories — Practical guidance for the estimation of measurement uncertainty

Laboratoires médicaux — Lignes directrices pratiques pour l'estimation de l'incertitude de mesure

This approach is often referred to as the IQC+EQA method

It principally captures (all) the components of a given test's contributing uncertainties by means of combining the calculated uncertainties derived from three components:

(1) the already assigned and stated MU value of the end-user calibrator (a calibrator or reference material with stated MU value) *However we do not have this for flow cytometry tests*

(2) the long-term precision (this captures random errors from reagents and measuring systems, over time), which is derived from IQC data, and finally

(3) the bias, which is derived from the laboratory's performance in EQA

The “top-down” approach

Or the IQC + EQA approach, also the LTUM

Internal Quality Control and External Quality Assessment (IQC + EQA)

$$1{,}96 \text{ x} \sqrt{u^2(IQC) + u^2(EQA)}$$

Long-Term Uncertainty in Measurement (LTUM)

$$1{,}96 \text{ x} \sqrt{LCV(a)^2 + LTB^2}$$

Calculations are reasonably complex:
Few Flow laboratories currently
- as far as I am aware-
provides a full IQC+EQA based MU,
Those who do, have statistician support.

Numerical MU at UHBW *at this time*

IQC data alone, a

For CD34 and Lymphocyte Subsets,

We use IQC (commercially available),

We calculated MU as precision; based on 3 separate lots. We determined that we normally performed within +/- 2SD. This was deemed acceptable.

Then, every month, we look at our lot for that month, and ensure we are still within +/- 2SD.

How does this improve safety/reduce risk?

The IQC gives us an idea of processing errors/analyst differences maybe

But only internally – we could create a very stable poor performance!

We do participate in EQA…and we aim to include EQA data in our MU

Calculating u IQC (IQC based uncertainty; precision)

According to the ISO Technical guidance, how to pool IQC lots and calculate uIQC

Calculation of Individual Lot's uncertainty for long term presicion (u_{IQC})

Data collection period:	03/04/24 - 09/05/24		
	Run	Low control	High control
Lot	(n=1/assay)	cells/ul	cells/ul
BC0424	1	9.8	29.0
BC0424	2	13.1	32.6
BC0424	3	10.4	31.6
BC0424	4	12.5	32.8
BC0424	5	11.5	34.3
BC0424	6	13.4	27.7
BC0424	7	14.6	30.0
BC0424	8	13.0	32.7
BC0424	9	13.9	33.0
BC0424	10	13.5	34.1
BC0424	11	12.5	28.9
BC0424	12	12.7	28.6
BC0424	13	11.3	34.3
BC0424	14	12.6	33.5
BC0424	15	14.3	26.5
BC0424	16	12.6	30.1
BC0424	17	11.2	32.3
n	17		
√n	4.12		
Mean		12.5	31.3
SD		1.3	2.5
u=SD/√n		0.3	0.6
u^2		0.1	0.4

Data collection period:		30/01/24 - 29/02/24	
	Run	Low control	High control
Lot	(n=1/assay)	cells/ul	cells/ul
BC0224	1	11.6	40.2
BC0224	2	10.4	35.2
BC0224	3	11.3	34.7
BC0224	4	11.2	41.9
BC0224	5	11.5	42.8
BC0224	6	12.1	46.6
BC0224	7	10.2	44.8
BC0224	8	11.5	38.4
BC0224	9	12.0	47.2
BC0224	10	11.3	45.9
BC0224	11	11.9	36.1
BC0224	12	12.5	39.6
BC0224	13	10.4	43.8
BC0224	14	12.6	40.7
BC0224	15	9.9	43.5
BC0224	16	10.8	40.6
BC0224	17	12.0	38.3
n	17		
√n	4.12		
Mean		11.4	41.2
SD		0.8	3.9
u=SD/√n		0.2	0.9
u^2		0.0	0.9

Data collection period:		03/06/24 - 17/07/24	
	Run	Low control	High control
Lot	(n=1/assay)	cells/ul	cells/ul
BC0624	1	9.5	35.1
BC0624	2	10.2	33.2
BC0624	3	8.1	36.1
BC0624	4	8.4	36.0
BC0624	5	8.6	35.2
BC0624	6	8.2	36.7
BC0624	7	7.9	36.5
BC0624	8	8.1	31.4
BC0624	9	7.6	34.3
BC0624	10	10.2	33.3
BC0624	11	8.4	35.4
BC0624	12	8.1	28.1
BC0624	13	9.1	34.7
BC0624	14	7.6	41.0
BC0624	15	7.8	37.2
BC0624	16	8.0	34.2
BC0624	17	8.9	35.4
BC0624	18	6.8	33.8
BC0624	19	8.0	36.6
n	19		
√n	4.36		
Mean		8.4	35.0
SD		0.9	2.6
u=SD/√n		0.2	0.6
u^2		0.0	0.4

NHS University Hospitals Bristol and Weston NHS Foundation Trust

IQC

3 separate IQC lots used, and provided the checks pass, their uncertainties are combined.

Combining the uncertainties of the three lots into one single uncertainty

Check 1: The data sets are assumed to have normal distribution. Any overt outlier (none of these datasets had outliers) would be removed.

Check 2: Is there a marked difference between the individual lots' uncertainty? No. The uncertainty values for the three lots high and low controls are very similar: the difference is less than 0.5 cells/ul. Thus, there is no indication that the slight differences seen would have any impact on clinical decision points.

If no marked difference between the lot's uncertaties, proceed to calculate the combined uncertainty for the three lots' low and high controls:

Low control:	$u_{IQC\ Low}$ (lots 0424, 0224, 0624) =	$\sqrt{\{(u_{Lot0424}^2 + u_{Lot0224}^2 + u_{Lot0624}^2) / 3\}}$
	$u_{IQC\ Low}$ (lots 0424, 0224, 0624) =	$\sqrt{\{0.011419368 + 0.004194449 + 0.003890194)/3\}}$
	$u_{IQC\ Low}$ (lots 0424, 0224, 0624) =	0.39
High control:	$u_{IQC\ High}$ (lots 0424, 0224, 0624) =	$\sqrt{\{(u_{Lot0424}^2 + u_{Lot0224}^2 + u_{Lot0624}^2) / 3\}}$
	$u_{IQC\ High}$ (lots 0424, 0224, 0624) =	$\sqrt{\{0.03992653 + 0.097675533 + 0.03531365)/3\}}$
	$u_{IQC\ High}$ (lots 0424, 0224, 0624) =	1.16

Then, add the confidence level. Usually 95%, meaning **x1.96** x the uncertainty

1.96 x 1.16=2.27 cells/ul

NHS University Hospitals Bristol and Weston NHS Foundation Trust

ESCCA Working group on MU

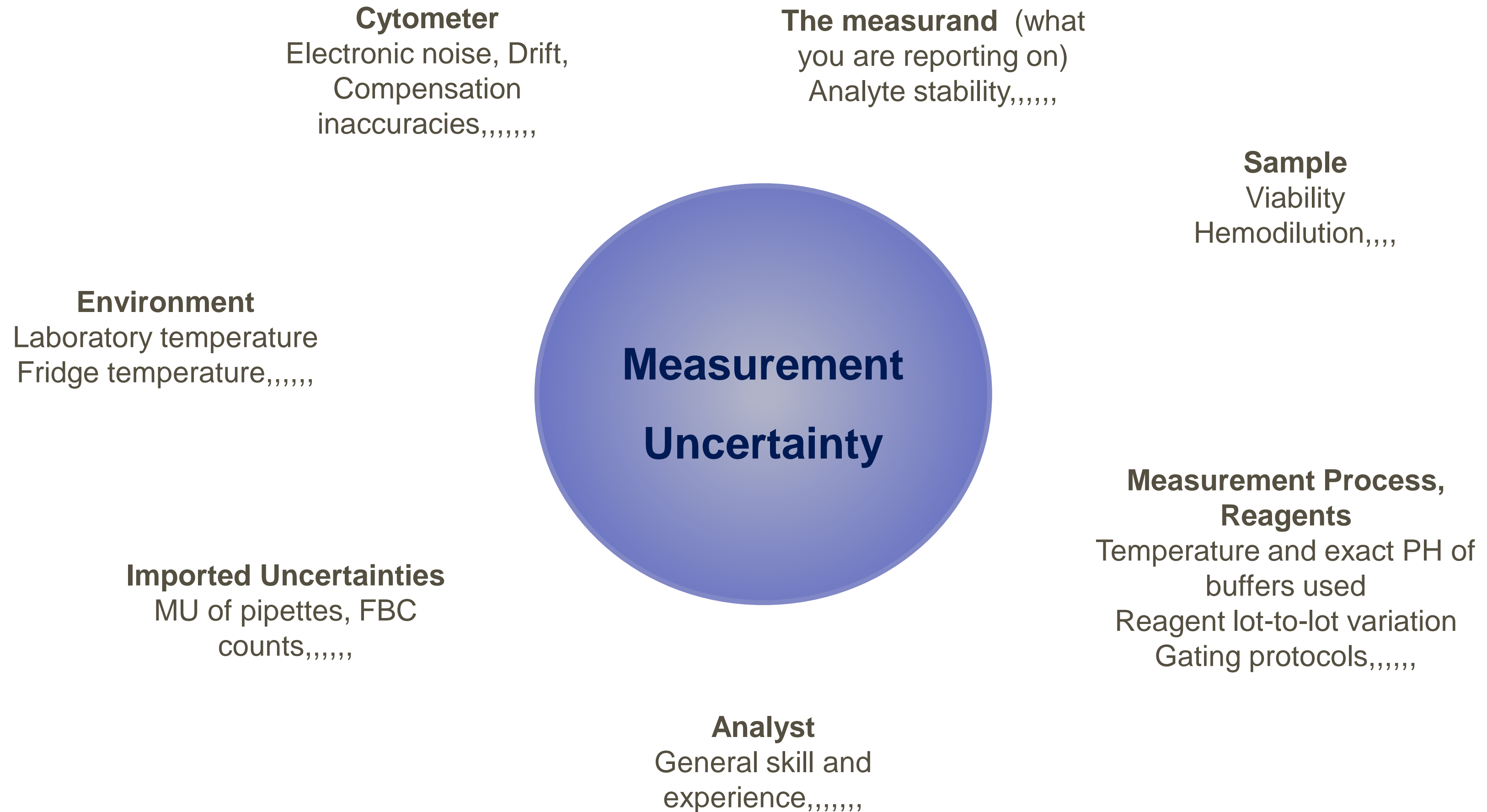
What we all very easily agreed on

" Importantly, the MU approach is meant to help to promote an overall high quality and a clinically safe flow cytometry service: not simply consume staff time for "tick-box purposes". The approach needs to be relevant to both the laboratory and the service users. It should help the laboratory to reduce variation as far as possible, as part of routine practice, to aid awareness of key contributing factors that could be well suited for audit, and ultimately whether it increases the overall robustness of the result and helps provide guidance for clinical interpretation of the test results. "

ESCCA recommendations for MU approaches in the clinical flow cytometry lab ***Manuscript submitted***

What can contribute to uncertainty?

Things that adds to high CV – or, worst case scenario – wrong result



Risk of generating wrong result

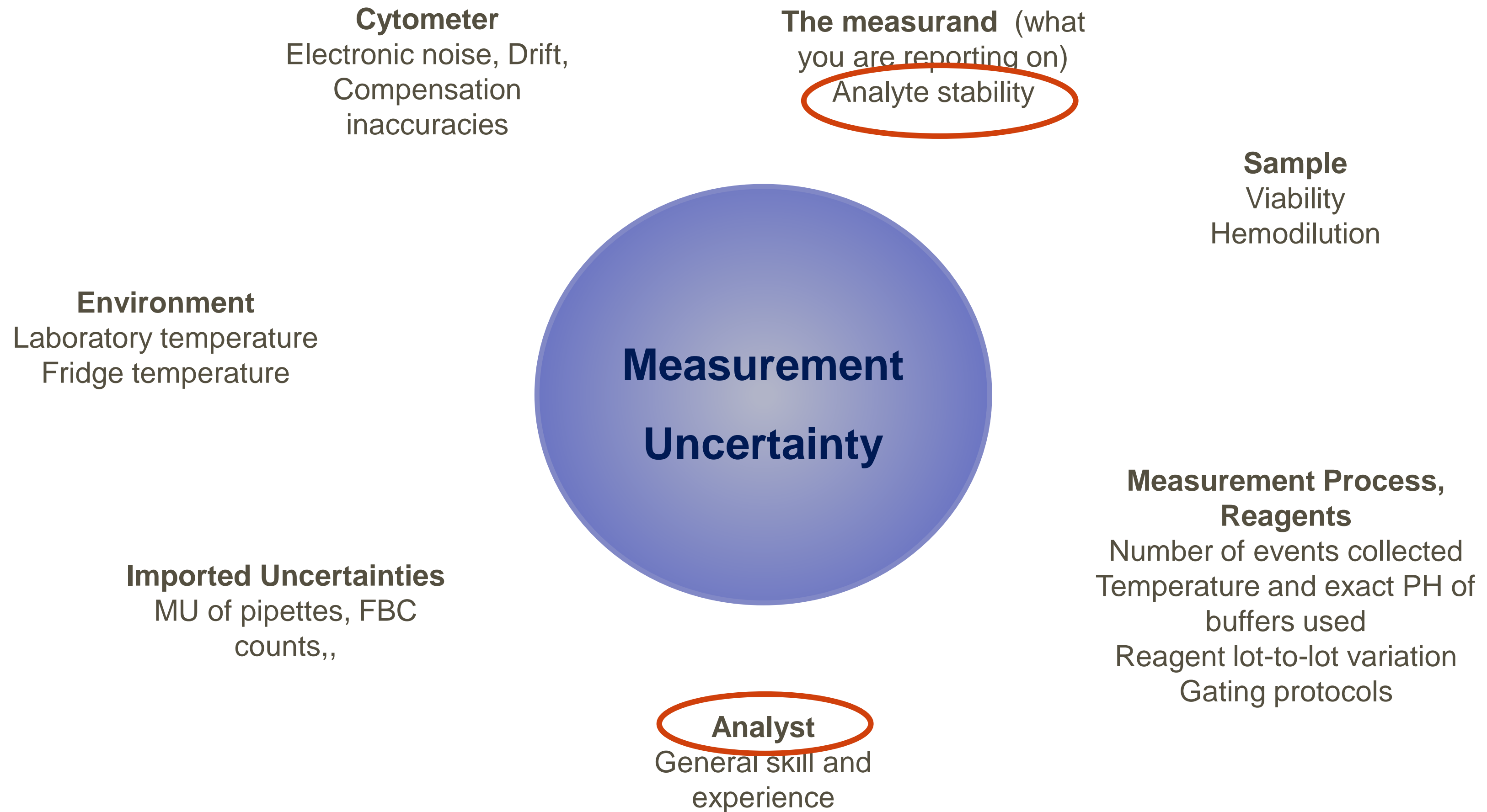
Similar to causes contributing to measurement uncertainty

Factors contributing to the uncertainty of a clinical flow assay result, their estimated relative contribution to uncertainty, and how their impact may be lessened, with an indication of the mitigated relative contribution to uncertainty.

Description	*Risk 0-5*	Critical to result (if so, describe why)	Mitigations in place to minimise uncertainty	*Mitigated risk 0-5*
Sample				
Sample volume: Insufficient	3	Particularly for MRD and other rare events assays. May result in failure to recover sufficient cells to reach the assay's desired LoQ and so contributing to overall uncertainty of the result. This includes CSF samples, especially during disease monitoring of CNS disease	Yes: Minimum volume set and communicated to service users. Training provided Samples with less than the stated minimum volume are processed where volume permits. If insufficient cells are available to reach a given assay's desired LoQ or LoD, the report is issued with a statement to that effect and, ifrelevant, the LLoQ for the sample may be stated.	1
Time of sampling: (1) delayed receipt of CSF samples (2) delayed receipts of non-urgent samples 	2-3	(1) If perishable samples such as CSFs arrive out of hours: cell death may occur and a falsely low cell count may be reported, and occult disease potentially may not be detected (2) Samples processed beyond validated analyte stability may produce inaccurate or suboptimal results . If the effect of sample age on assay read-out is unknown, the analyst will not recognise changes from norm that are due to or, could be due to, sample age rather than neoplasia or reactive causes.	(1) A call–out system is in place for the most urgent types of queries where sampling out of hours. Preservative is used for CSF samples (2) Sample age is stated on the sample itself and request form. If information is lacking, staff contact clinic to find out. Analyte stability is investigated as part of assay *validation process* Analysts therefore know what/which parameters may be reportable for fresh and older samples, and a cut-off may be used. Due to the nature of disease being investigated and invasiveness of sample taken, clinical indication may require sample to be tested outside the recommended time frame for the assay. If tested outside the recommended time frame, results should be reported with suitable comment.	1

What did contribute mostly to uncertainty?

Looking at precision data from Assay validation work:



Using analyst precision on a reasonably regular basis (EQA samples for example) to have idea of variability in results reported form the lab.

The below example uses data from an EILCP Chronic Lymphocytic Leukaemia Measurable Residual Disease exercise.

	% MRD (of total cells)	
Participant	Original	High outlier removed
1	0.20	0.20
2	0.18	0.18
3	0.16	0.16
4	0.18	0.18
5	0.19	0.19
6	0.14	0.14
7	0.18	0.18
8	0.17	0.17
9	1.07	
10	0.17	0.17
11	0.19	0.19
12	0.19	0.19
13	0.19	0.19
14	0.18	0.18
15	0.16	0.16
16	0.17	0.17
17	0.17	0.17
18	0.14	0.14
19	0.18	0.18
20	0.18	0.18
21	0.17	0.17
22	0.18	0.18
23	0.20	0.20
24	0.17	0.17
25	0.19	0.19
26	0.18	0.18
27	0.18	0.18
n	27	26
√n	5.196	5.099
Mean	0.210	0.177
SD	0.173	0.015
u	0.033	0.003
k	2	2
Expanded uncertainty (U)	0.066	0.006
Lower 95% CI of mean	0.141	0.171
Upper 95% CI of mean	0.278	0.183
% U_{rel}	16%	1.7%
CV	82.3%	8.5%

EILCP

We are 3 analysts , and use EQA (NEQAS, EILCP) to compare our results,

Some labs refer to this as MU (we don’t but t

There is some support in ISO guidance on MU for this approach.

During assay validation – the analysis template often re-made to reduce discordance between analysts.

-> Are some assays higher risk than others – where should we place our efforts??

Risks…

The current UHBW approach

In broad, risks in the flow cytometry lab may relate to:

- Health and safety for staff or visitors.
- Risk of ‘major events’.

  Emergencies including electricity outage, fire, prolonged cytometer failure/breakdown, severe staff shortage due to e.g. pandemic.
- Risk of generating incorrect results.
- Risk to patient should wrong (or no) results be reported

Risk to patient (if wrong/no result is provided)

What other results/clinical information is available, to mitigate impact of false flow result?

		CD34	Ly subset	Sezary Diagnosis / Monitoring	PNH
Risk to patient (describition)	Falsely elevated count/false positive (wrong description of population of interest)	Insufficient harvest time=> potential need for another harvest session (which may be difficult to arrange if no slots available). Potentially also missing peak harvest time, risking overall insufficient collection.	Result is used a guidance and not for harvest decision	At diagnosis: Blood staging will raisee from B0 to B1 if count is > 250 cells µL−1; and from B1 to B2 if count is >1000 cells µL−1. Staging takes several factors into account and a high (>1000) sezary count is visible on morphology, thus film review helps mitigate result. Monitoring: may influence therapy desicion however overal picture takin into account and if results not matching pictute, a repeat sample may be sent.	Diagnosis: waste of time, confusion: overall picture unlikely, after ocnsideration, match the flow result - a repeat flow may be sent. Monitoring: confusion, if result does not match other counts, and expected response, a repeat smaple is liekly requested.
	Reduced count / abnormal/disease population missed	Harvest not taking place, missing peak harvest time, risking insufficeint collection and need for another round of G-CSF/Plerixafor , wait and harvest.		Diagnosis: If CTLC was present in circulation, but missed by flow test, this may redner diagnosis more diffiult, and may cause a delay, due to the not always easy diagnosis procedure for these disorders. Monitoring: missed disease, if sizeable, may affect physician/MDT 's therapy decision, th eoveral lclinical picture is taken into account.	Diagnosis: if a clinically relevant PNH clone is missed, correct therapy may be delayed, thrombotic or haemorragic event may occur, that could have been prevented.
Risk Rating	Consequence:inherent	Moderate	Negligeble	Moderate	Moderate/Major (for frank PNH)
	Likelihood: inherent	Possible	Possible	Possible	Possible
	Risk rating: **Inherent**	9	3	9	9 to 12
	Consequence:with mitigations	Moderate	Negligeble	Moderate	Moderate/Major
	Likelihood: with mitigations	Rare	Rare	Rare	Rare
	Risk rating: W**ith mitigations**	3	1	3	3 to 4
Mitigations	Mitigating surrounding clinicopatholgical data	Yes: FBC count, stem cell count post harvest	Yes: FBC count, stem cell count post harvest	FBC count, film review, skin biopsy	Frank PNH: clinical symptoms, LDH, other haemolyis-related infromation, film review; in context of AA/MDS: very low risk.
	Integrated report	No	No	Yes	No
	MDT	No	No	Yes	Yes

Thank you!

Questions?

Catarina Martins
Eleni Liskens
Evi Yiannaki
Kathryn Clarke
Neil McIver-Brown
Richard Veyrat-Masson
Thomas Keller
Dan Payne
Dave Bloxham
Francesco Buccisano
Juan Flores-Montero
Katherina Psarra
Katy Jones
Maria João Acosta
Ulrich Sach
Ulrika Johansson
Wolfgang Kern

Title

Subtitle

Text

Reference