# Characterising performance of *in situ* hybridisation/tissue methodologies in order to demonstrate fit-for-purpose assays for clinical deployment

**Neil Henderson**
**Personalised Healthcare & Biomarkers,**
**AstraZeneca**

AstraZeneca

# Aims

## The challenge

- Challenged with establishing a PoM protocol for evaluating biomarker x in human biopsy material.


- Test the hypothesis that the levels of biomarker x (which according to literature was elevated in disease) would change after treatment with drug.

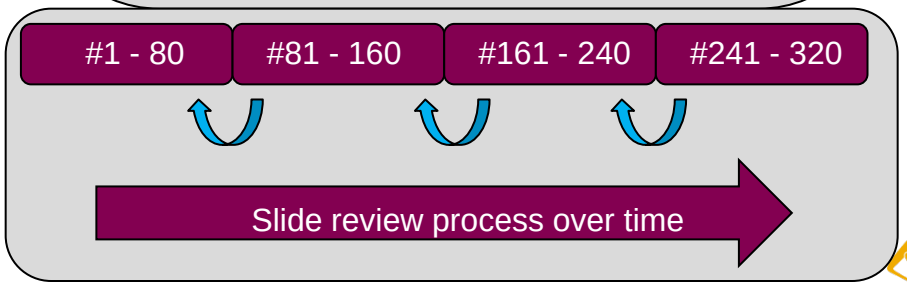
# Basics of tissue analyses

## IHC & ISH

IHC – immunohistochemistry (protein)
ISH – in situ hybridisation (mRNA)

- ISH/IHC are techniques for qualitatively/semi-qualitatively assessing the distribution of a biomarker within a tissue section
- The 3-5 micron sections are taken from a wax or frozen block. Potentially each section could be distinct
- These techniques are labour intensive & time sensitive
- There are a lack of positive controls
- Negative (non-specific) staining can occur

- Scoring involves trained pathologists
- Assessment of stained sections can incorporate a number of factors: subjective or partially quantitative
- Scoring system can be simple or complex
  - Amount/distribution of stain
  - Number of cells stained
- Scoring system used
  - 0 – 5 where 0 = none to 5 = severe
- Pathologist must constantly review slides in order to maintain consistency



# Validating IHC/ISH guidance

- **LBA/MS**
- **Established guidance**

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Guideline on bioanalytical method validation

MHRA

Good Clinical Practice

- **IHC/ISH**
- **Draft FDA guidance document 2009**

**Outcome/aim: interpret the relevant guidance documents in order to demonstrate a reliable (fit for purpose) IHC/ISH method**

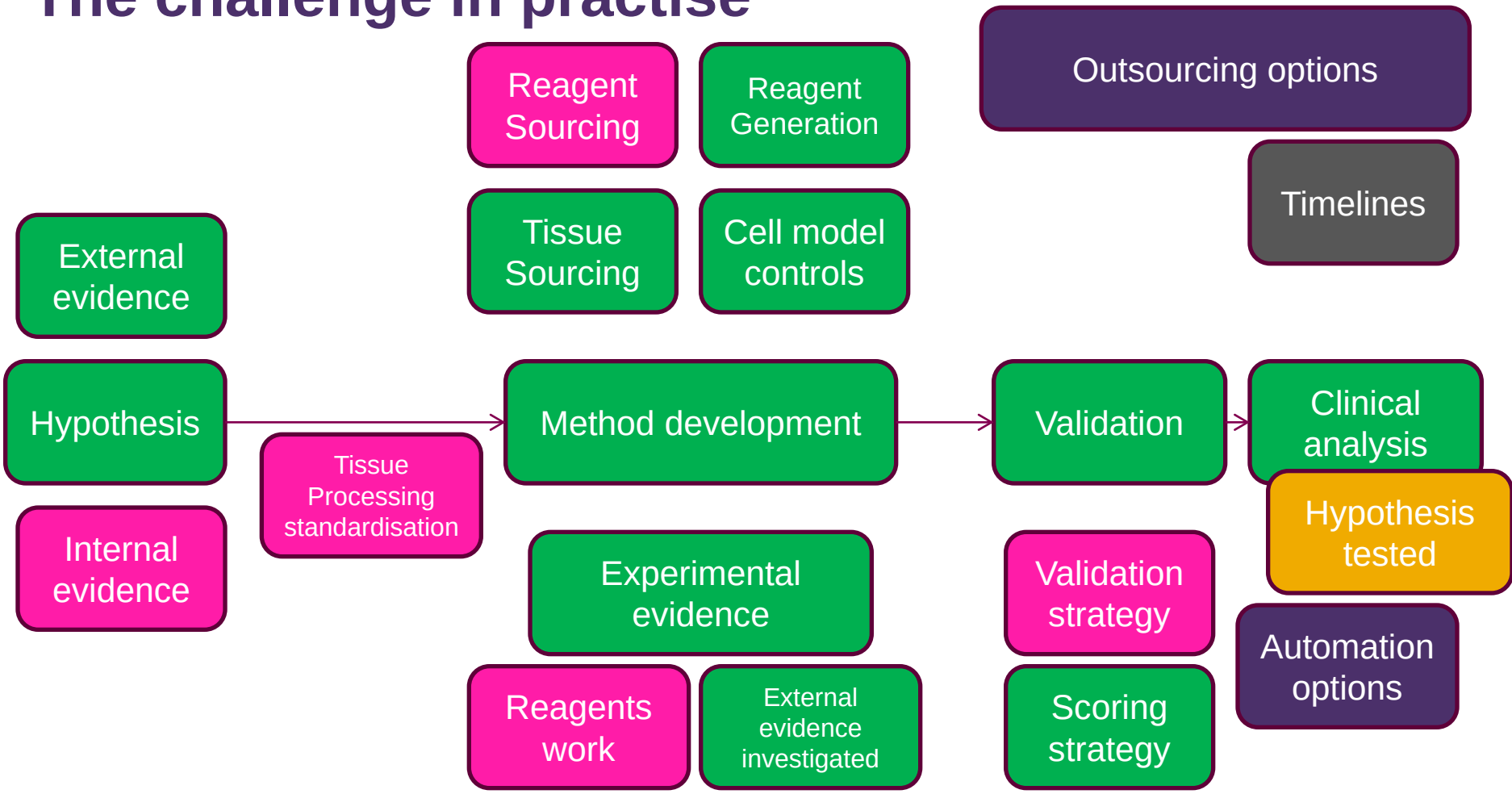
# Aims

## The challenge

- Simple plan

# Aims

## The challenge in practise



# The journey begins – 1st stop: sample collection

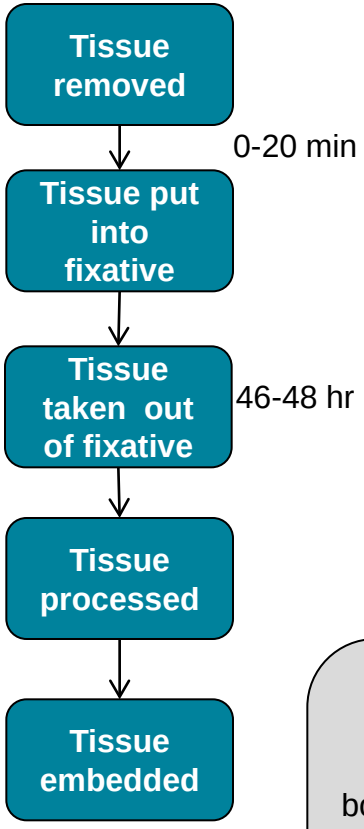
## Ensure optimum sample collected

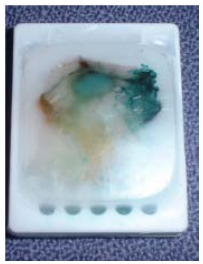
- Sample collection is one of the most important factors for any biomarker analysis
- Blood methods – universal blood collection tubes; avoidance of haemolysis etc

**Tissue collection – several areas for generating an inconsistent sample**

- Size of biopsy collected (affects fixation process characteristics)
- Time out of patient into formalin (sample still biologically active)
- Time in formalin (over/under fixation)
- Length of dehydrating tissue processing step (over/under processing for biomarker
- Orientation of biopsy in the embedding cassette (can affect what is being measured
- This was a MULTI-CENTRE site collection

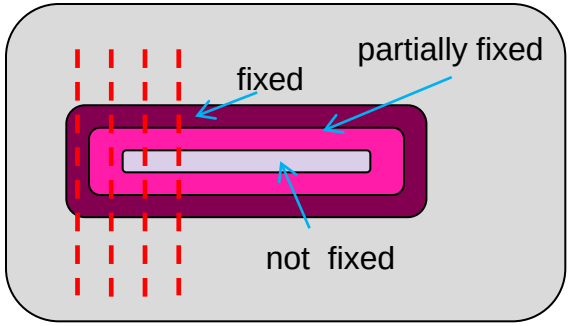
**Outcome: Specified consistent collection regime**

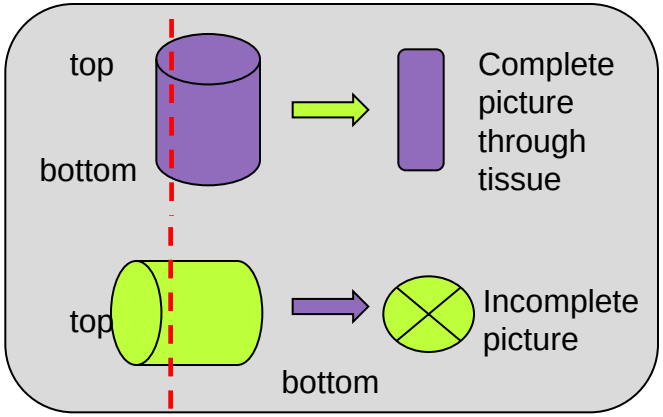


Large Biopsy

Needle biopsy




Biopsy orientation

# Finding suitable reagents

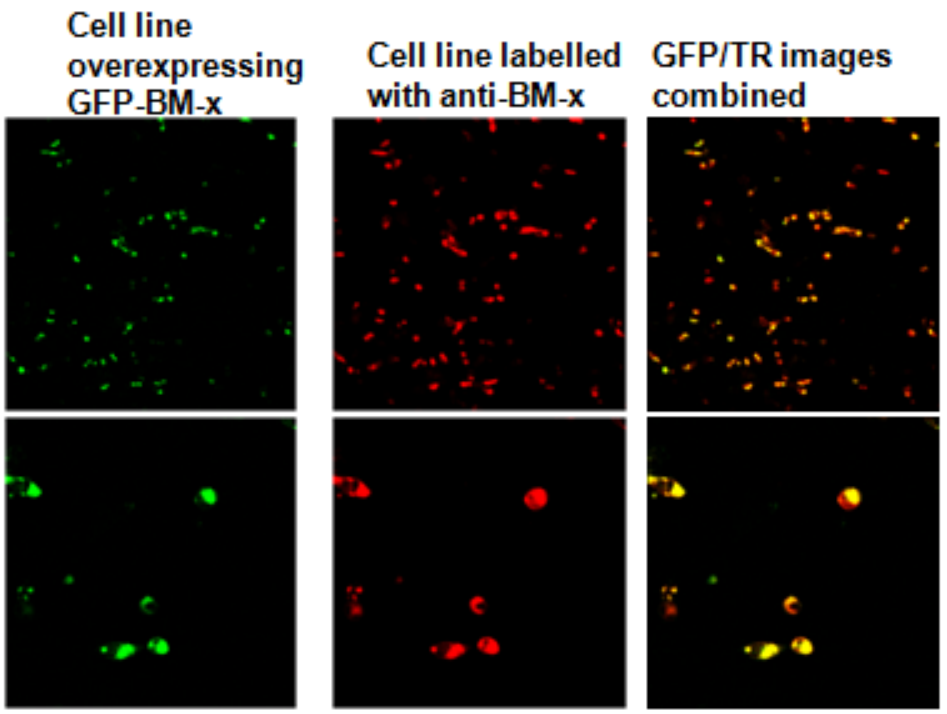
## Demonstrating specificity in cell models

- IHC requires specific antibodies to the target
- Some antibodies bind to target in cell models but not to target in tissue

- Primary Task: source commercial or generate novel antibodies
- Source cell control models to show target specificity
  - Immunocytochemistry (ICC)
  - Western blot analysis
  - GFP-overexpression models
  - MS signatures

Outcome: the data generated gave confidence that the antibodies were selective

- The entire lot number/batches of commercial antibodies were purchased to avoid having to perform lot to lot variability studies



# Reagent use in tissue

## Transferring utility

- Confident that abs are specific

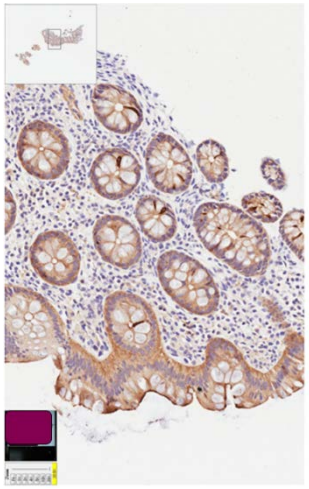
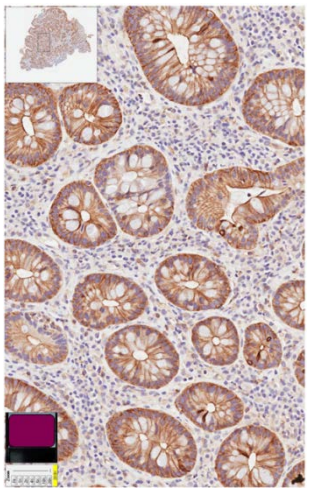
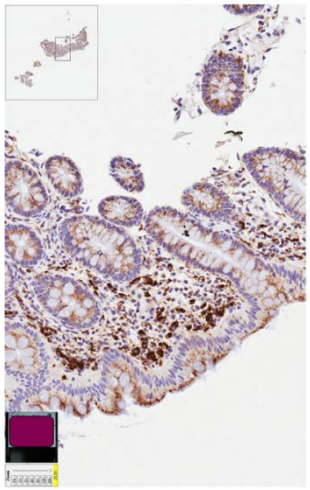
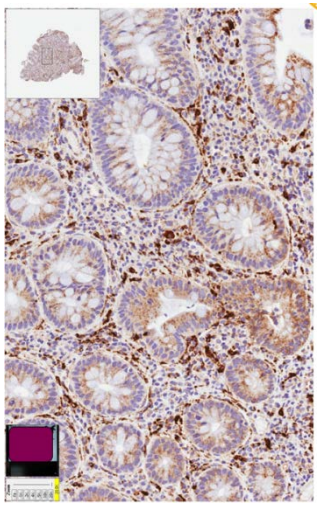
**Task: do they work in tissue?**

- BM-Y
  - Shows specific binding (where expected to have binding)
  - No differentiation in stain intensity between non-diseased/diseased
  - Signal optimised (but not suitable for criteria)

**Outcome: BM-Y dropped**

- BM-X
  - Stain intensity appeared to be correlated to disease severity
  - Unexpected binding pattern

**Outcome: ISH (mRNA signal) investigated as a support methodology for IHC**

| | Non-diseased | Diseased |
|---|---|---|
| BM-Y |  |  |
| BM-X |  |  |

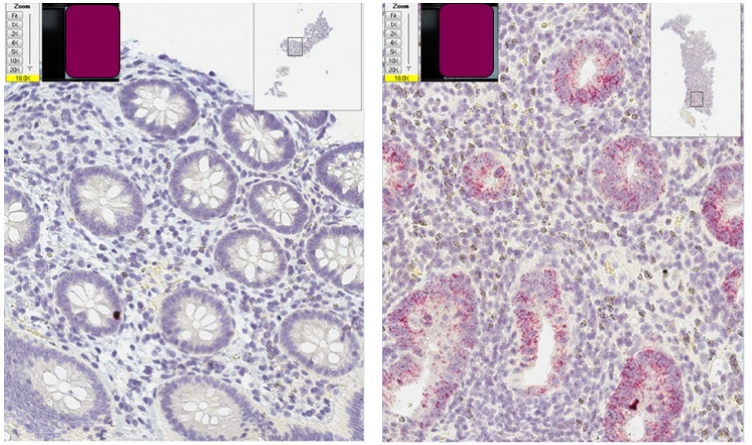
# Reagents in tissue (pt2)

## Evolving strategies

- ISH uses a cDNA probe to target mRNA
- Rationale: run both IHC & ISH – overlay images to allow pathologists eye to identify true BM-X signal
- BM-X ISH
  - Low background
  - Signal correlated to disease severity
  - Possibility for automation
  - Corroborated IHC staining

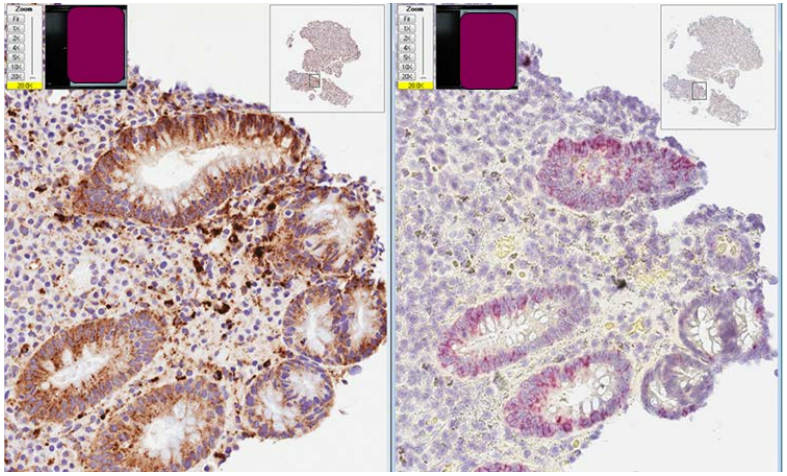
Outcome: BM-X ISH chosen to be taken forward as potential PoM biomarker assay in its own right

ISH



Non-diseased Diseased

Diseased

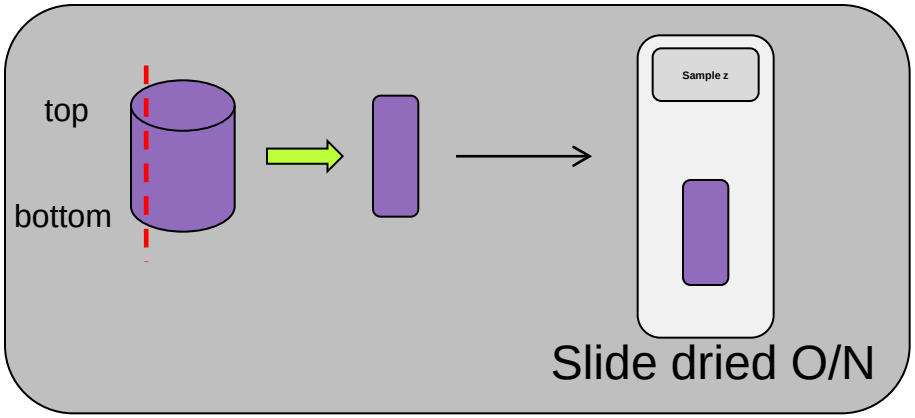


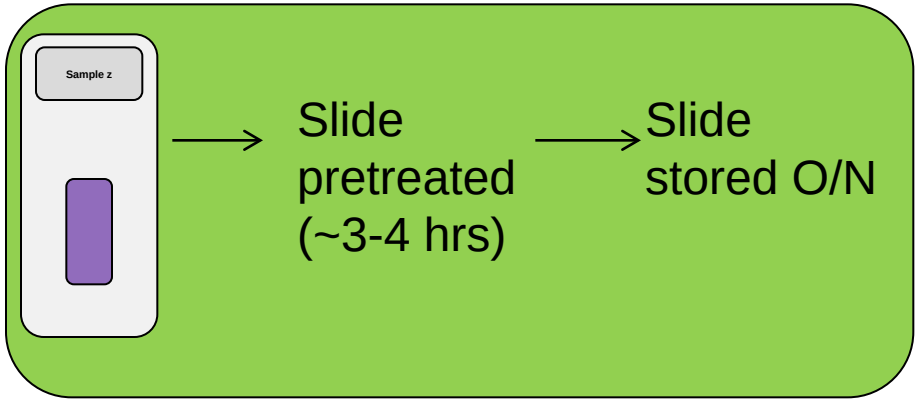
IHC ISH

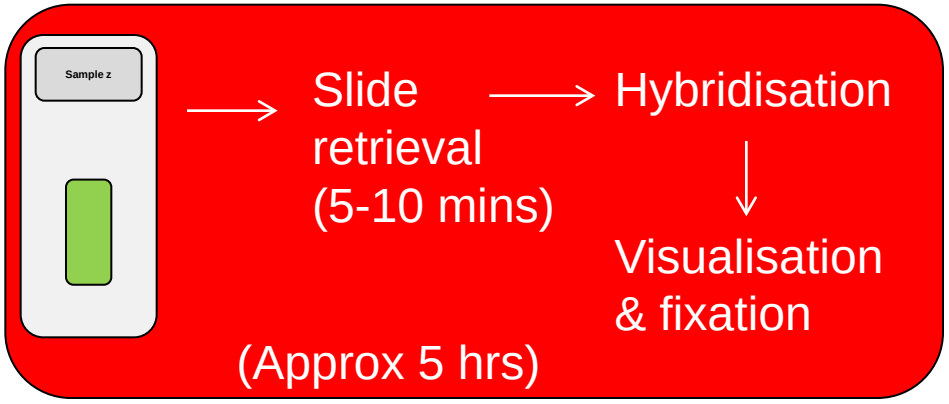
# ISH

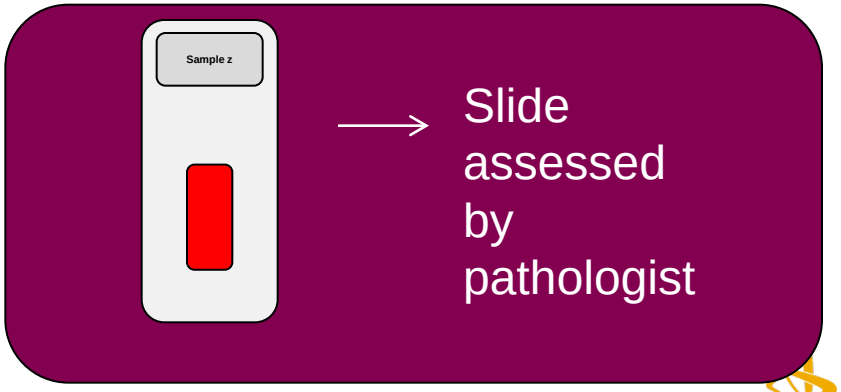
## Methodology in detail

- By hand method (20 slides per batch possible)



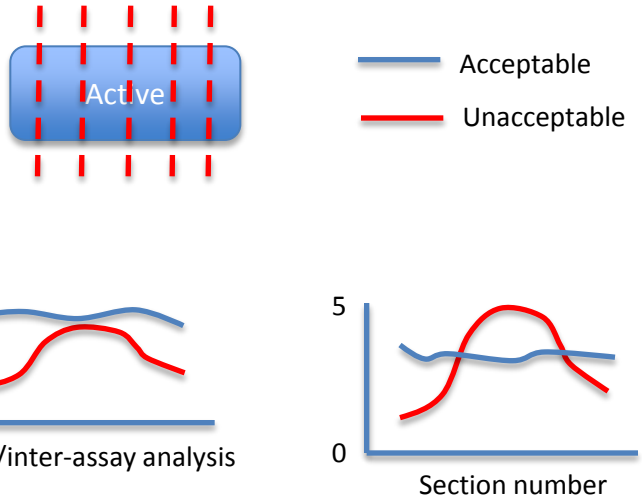



# Validating tissue assay: pre-concerns I

## Cutting through a biopsy

- Biopsies are heterogeneous
- Samples can not be made homogeneous
- When taking a section for analysis – it is potentially different from the last section taken
- How are you supposed to demonstrate reproducibility of an assay when there is chance of no fixed point for reference
- Potential for different result each time is high
- Variability between samples is potentially high
- Reading window is potentially too small

**Potential variability within a sample**




Reproducibility of the performance of the assay must be shown on the same day (intra) & over several days (inter) for the assay to have clinical robustness

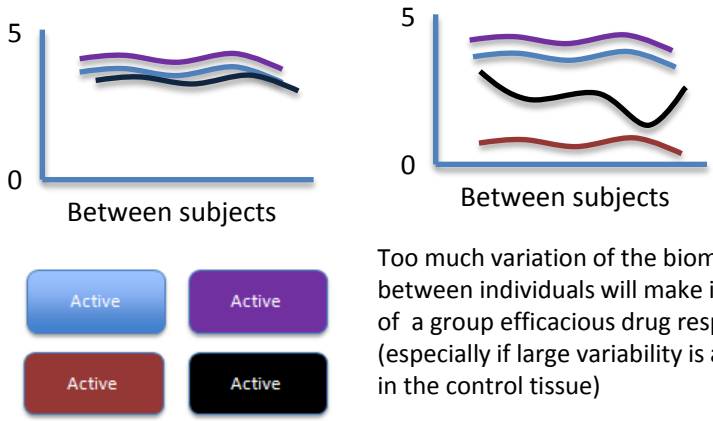
If biomarker expression is variable throughout a tissue block, then a greater number of sections potentially would need to be analysed to give assay confidence (but would lead to increased cost $$)

# Validating tissue assay: pre-concerns II

## Cutting through a biopsy

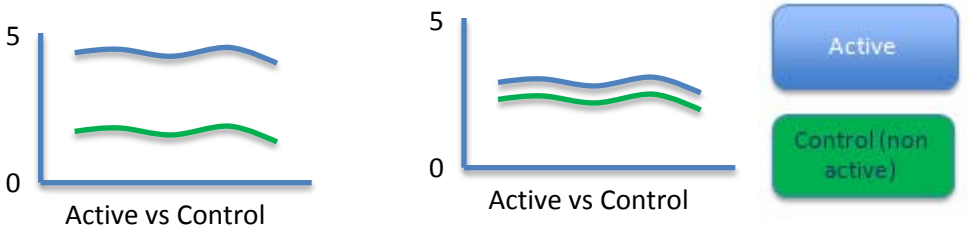
- Biopsies are heterogeneous
- Samples can not be made homogeneous
- When taking a section for analysis – it is potentially different from the last section taken
- How are you supposed to demonstrate reproducibility of an assay when there is chance of no fixed point for reference
- Potential for different result each time is high
- Variability between samples is potentially high
- Reading window is potentially too small

### Variability between different samples




Too much variation of the biomarker between individuals will make interpretation of a group efficacious drug response difficult (especially if large variability is also reflected in the control tissue)

### Differences between Active vs Control samples




If "reading window" is too narrow then demonstration of an efficacious drug response will be difficult to show & interpret

# Validation strategy

## ISH

### Cell pellets (embedded in paraffin)

- Positive control (over expressing BM-X)
- Negative control (no expression of BM-X)

- Sample regarded as “more homogeneous”
  - Inconsistencies with analytical process would be highlighted

- Intra- & inter-assay variability
- Between analyst variability (3 analysts)
- Pre-incubation stability

### Tissue biopsies

- Non-active diseased (baseline)
- Active diseased (range of biological expression)

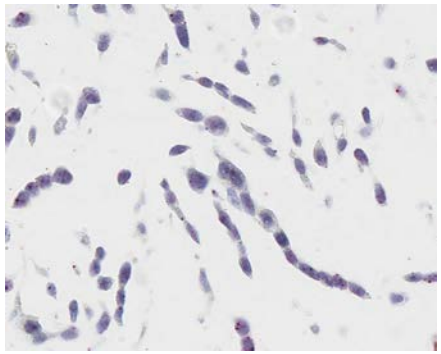
- Samples regarded as “potentially highly heterogeneous”

- Distribution (range)
- Variability between adjacent sections
- Variability between non-adjacent sections
- Variability between independent biopsies from same individual

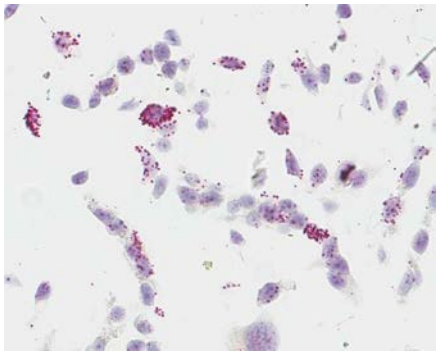
# Validation

## Intra-assay & inter-assay reproducibility

- Cell blocks used
  - BM-x probe (positive signal)
  - Sense (negative signal)



SENSE probe

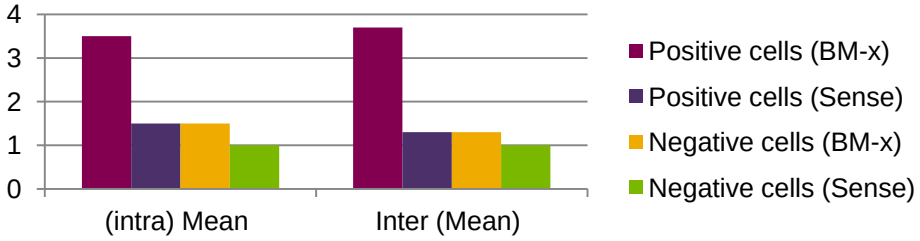


Biomarker X probe

- Amount of red (positive) stain scored 0 – 5

Outcome: Intra & Inter assay reproducibility demonstrated
Data reported as Mean only

| sample | Positive cells (BM-x) | Positive cells (Sense) | Negative cells (BM-x) | Negative cells (Sense) |
|---|---|---|---|---|
| 1 | 3 | 1 | 2 | 1 |
| 2 | 3 | 2 | 1 | 1 |
| 3 | 4 | - | 2 | - |
| 4 | 4 | - | 1 | - |
| (intra) Mean | 3.5 | 1.5 | 1.5 | 1 |
| 5 | 4 | 1 | 1 | 1 |
| 6 | 4 | 1 | 1 | 1 |
| Mean | 4 | 1 | 1 | 1 |
| Inter (Mean) | 3.7 | 1.3 | 1.3 | 1 |

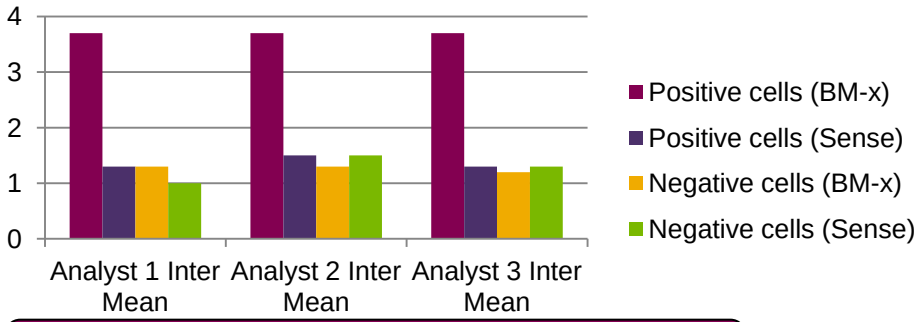



Data from primary analyst

# Validation

## Inter-analyst reproducibility

- 3 independent analysts used
- Needed to demonstrate the degree of analyst influence on the staining of the cells




Outcome: comparable data generated

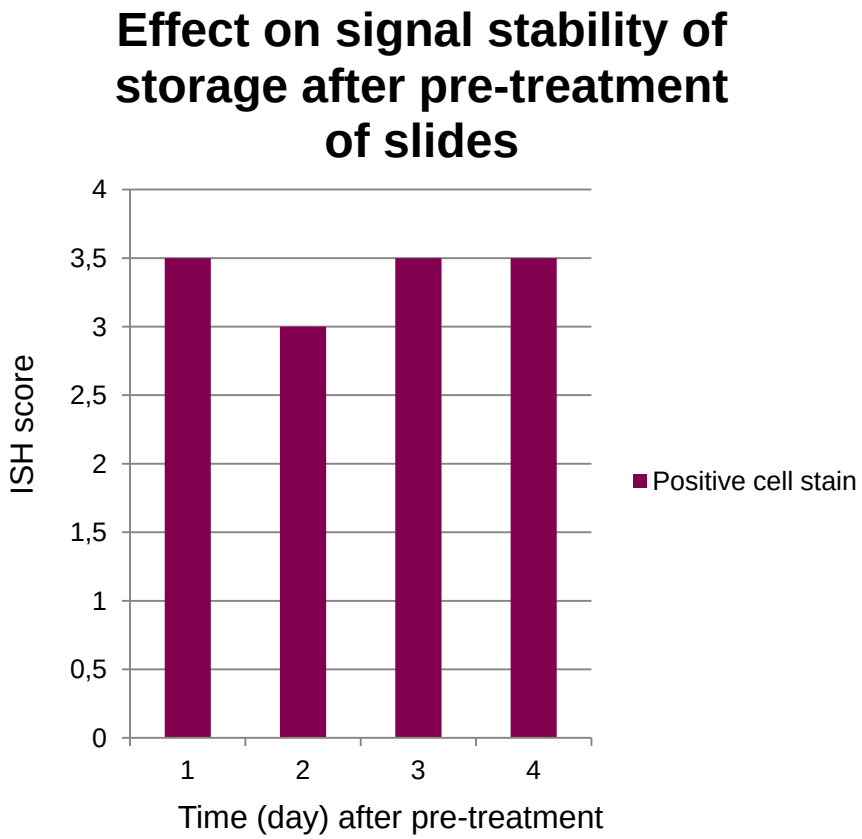
| sample | Positive cells (BM-x) | Positive cells (Sense) | Negative cells (BM-x) | Negative cells (Sense) |
|---|---|---|---|---|
| 1 | 3 | 1 | 2 | 1 |
| 2 | 3 | 2 | 1 | 1 |
| 3 | 4 | - | 2 | - |
| 4 | 4 | - | 1 | - |
| 5 | 4 | 1 | 1 | 1 |
| 6 | 4 | 1 | 1 | 1 |
| 7 | 4 | 2 | 2 | 2 |
| 8 | 4 | 2 | 1 | 2 |
| 9 | 3 | - | 2 | - |
| 10 | 4 | - | 1 | - |
| 11 | 4 | 1 | 1 | 1 |
| 12 | 3 | 1 | 1 | 1 |
| 13 | 4 | 1 | 1 | 1 |
| 14 | 3 | 1 | 1 | 2 |
| 15 | 3 | - | 1 | - |
| 16 | 4 | - | 1 | - |
| 17 | 4 | 2 | 1 | 1 |
| 18 | 4 | 1 | 2 | 1 |
| Mean | 3.7 | 1.3 | 1.3 | 1.3 |
| Range | 3 to 4 | 1 to 2 | 1 to 2 | 1 to 2 |

# Validation

## Slide preparation stability

- Prior to analysing cut sections for a biomarker of interest (by ISH), the sections must undergo a pre-treatment regime to “reveal” the cDNA target
- Normal practise was to analyse one day after pretreatment
- However, we wanted to understand if the samples could be left for a longer period post “pre-treatment”.

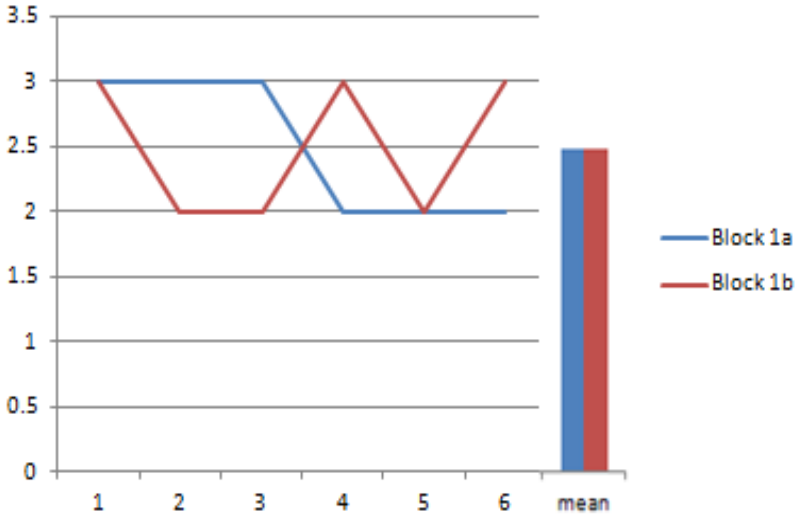
Outcome: no observed deterioration in signal intensity up to 4 days after pre-treatment



# Validation

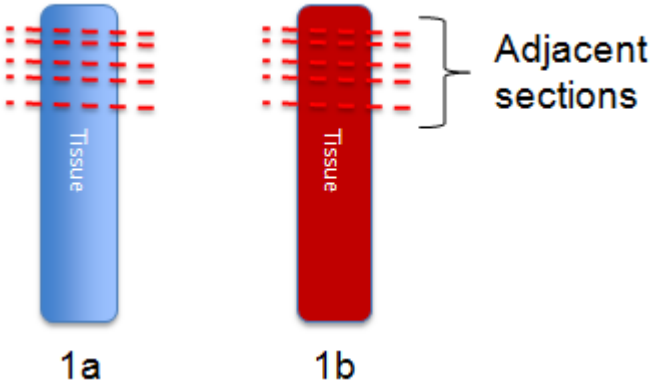
## Signal in tissue (adjacent sections)/duplicate biopsies

- Distribution of biomarker in the novel sections
- What is the variability
  - Between sequential sections
  - Between 2 biopsies taken from same individual (at same time)
- Investigated these factors to generate confidence that a signal detected when assay is clinically deployed isn't "random".




Outcome: the variability of signal appears conserved.

Outcome: Consistency between biopsies indicating that a change at post dose (from pre dose) could have clinical relevance

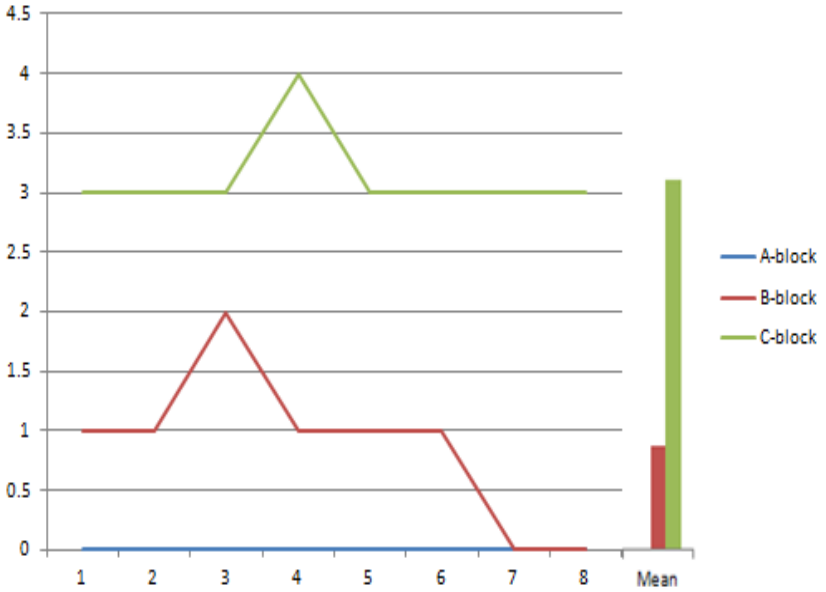


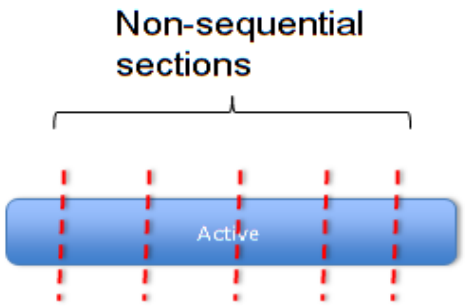
# Validation

## Signal in tissue (non-adjacent sections)

- Distribution of biomarker in the novel sections
- What is the variability
  - Between non-sequential sections
    - None (a-block)
    - Low (b-block)
    - Mid-Severe (c-block)
- These are small biopsies, so is there a finite amount of reproducible analyses that can be performed?

Outcome: the data indicates that a high degree of consistency occurs through most of the biopsy block



# Validation

# Conclusions

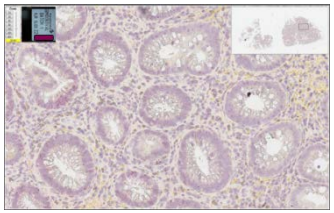
- Methodology shown to be reliable
- Use of multiple analysts shows consistent results
- No deterioration occurs in the stain intensity after storing pre-treated slides for up to 4 days
- Variability through a biopsy block shown to be relatively limited

Outcome: assay ready for clinical deployment

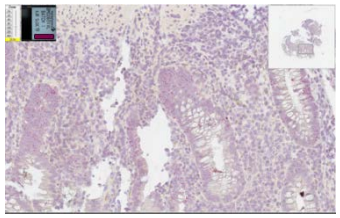
# Clinical deployment

## The real test for the assay

- 18 samples
- 2 cell controls (sense & BM-x probe)
- Only paired samples at pre & post dose were analysed (81% samples collected)
- Scored by pathologist (from validation)



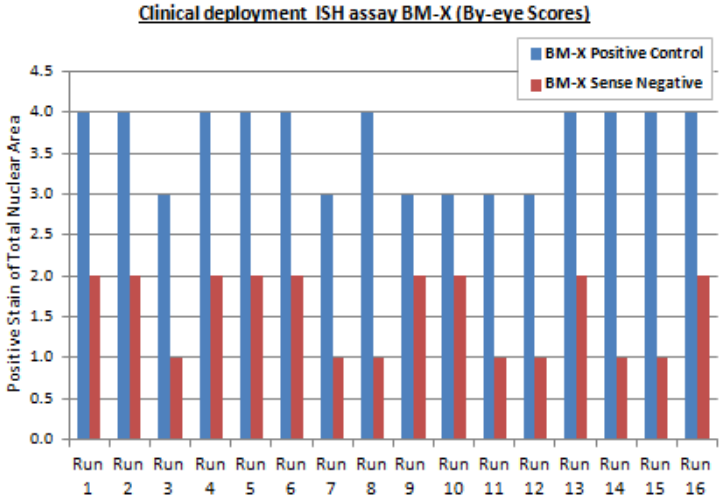
None/Low



Mid/severe

- Batches passed depending upon controls
  - BM-x stain > SENSE probe (Yes/No)
- All runs passed
- Data supplied on time for study database lock

Outcome: delta changes from pre to post dose observed




Delta change (pre to post)

| 5 | 4 | 3 | 2 | 1 | 0 | -1 | -2 | -3 | -4 | -5 | | NR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 8 | 5 | 27 | 57 | 23 | 13 | 5 | 1 | 0 | | 2 |

| Increase | None | Decrease | | NR |
|---|---|---|---|---|
| 42 | 57 | 42 | | 2 |

# Tissue analyses

## Value of work & the future

### Value of work

- Team established a reliable method and executed it to provide data within timelines
  - Sample analysis (wet work) was completed in 23 days
  - Pathologist read was completed in 25 days
  - Clinical analysis completed in 38 days

### Future

- For studies in the future the aim would be to increase the level of automation of the methods e.g. Ventana platform
  - Potentially greater analytical consistency
  - Reduce analytical pressure on analyst
- Evaluate computer algorithm to quantify the amount of staining
  - Correlates with pathologists subjective score
  - Computer algorithm could give a percentage of stained area vs total area
  - Computer algorithm would need to be validated and tissue sampling un-biased



# Conclusions

## Reflections

- “By hand”, multiple analyst tissue based analyses on biopsy material can reproducibly generate clinically relevant data for projects
- Cell pellets are suitable for pass/fail acceptance criteria test samples
- In the validation, the data indicated increasing (biopsy) section number did not notably change the outcome of analysis

- Essentials for success:
  - Establish consistency for sample collection & preparation
  - Establish confidence in reagents (cells & tissue)
  - Keep things simple
  - Ensure you have a great team of people to rely upon

# Acknowledgements

## A great team


- *Discovery Sciences, Alderley Park*
- Sonia Houghton

- *DSM Pathology Sciences, Alderley Park*
- Sally Price
- Kevin Randall
- Jaimini Reens
- Susan Smith
- Melanie Galvin

- *DSM Pathology Sciences, Mölndal*
- Gerhard Böttcher

- *CVGI Translational Sciences, Mölndal*
- Ann-Christin Nyström

- *PHB, Mölndal & Alderley Park*
- Clare Balendran
- Glen Hughes

- *Biobank, Mölndal*
- Ann-Sofi Andersson

- *Biobank, Alderley Park*
- Douglas McKechnie