


Defensible Decisions in an Uncertain World




Concepts

- Bulk vs attribute sampling
- Sampling is greatest source of error
- Analytical quality control and analytical error measurement is important for data acceptance, but what about sampling quality control and sampling error?
- Sampling starts with objectives
- Representative sampling includes mass, tools, increments, randomness, integrity

FSMA

.
"(6) Model laboratory standards.--The Secretary shall develop model standards that a laboratory shall meet to be accredited by a recognized accreditation body for a specified sampling or analytical testing methodology and included in the registry provided for under paragraph (1). In developing the model standards, the Secretary shall consult existing standards for guidance. The model standards shall include--
"(A) methods to ensure that--
"(i) appropriate sampling, analytical procedures (including rapid analytical procedures), and commercially available techniques are followed and reports of analyses are certified as true and accurate;
"(ii) internal quality systems are established and maintained;
"(iii) procedures exist to evaluate and respond promptly to complaints regarding analyses and other activities for which the laboratory is accredited; and
"(iv) individuals who conduct the sampling and analyses are qualified by training and experience to do so; and
"(B) any other criteria determined

What is Appropriate?

- "(A) Methods to ensure...true and accurate"
- What does this mean?
- Only samples that meet project objectives with a known or measured confidence can achieve this
- How is this possible?
  - develop objectives (SQC)
  - develop protocol to meet objectives (TOS)
  - assess that objectives were met (DQA)

Analyses for Decisions

- Are at the heart of everything we do
- Correct decisions are critical to our programs
- Decisions based on analytical results from field samples can be incorrect because of
  - analytical ***errors***
  - laboratory sample preparation ***errors***
  - primary sample collection ***errors***
  - ***errors*** in interpretation of results (including QC)

Decisions

- How does error cause incorrect decisions (decision error)?
- What can be done to reduce sampling error and thus decision error?
- How is sampling error measured?
- How to allocate resources based on tolerable decision error?

Example of Decisions

and Decision Error

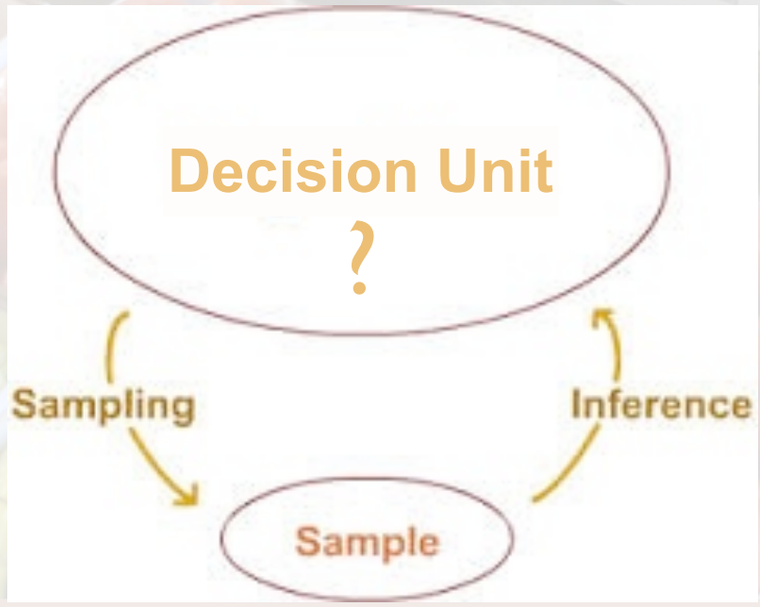
- Are there any blue beads?
  - take three beads out of Jar
  - take ten beads out of Jar
- What is the red bead concentration?
  - take three marbles out of Jar
    - no red ones (underestimate)
    - one red one (overestimate)
    - which is more likely?
  - take ten marbles out of Jar
- Do we make better decisions from three or ten marbles?

Sampling

- Should the marbles be selected in any particular way?
- If the marbles are different sizes and densities
  - where will the dense marbles be?
  - where will the fine marbles be?
  - where will the light marbles be?
  - where will the larger marbles be?

Sampling for Red Bead Concentration

- Is a single bead representative of all the beads? Then why collect it? Why analyze it?
- If a representative sample of beads was collected, would we have sampling error?
- What makes a primary sample representative?
  - some of everything in same proportion
    - mass
    - increments
    - tools
  - maintain analyte integrity



Decision Error

- If the analyses were perfect...
- If the sample processing was perfect...
- If the field samples were perfect...
- If the objectives were correct...
- If the data interpretation was correct...
- If you won the lottery last night...

Comes from sampling, sample processing, and analytical error

5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5

Mean = 5, Error = 0

5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2
5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2
5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2
5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2
5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2
5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2
5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2
5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2
5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2
5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2	5+/-2

Mean = 5, Error = 35%

7	6	5	6	1	6	4	5	4	4
5	3	8	2	3	4	6	8	6	4
6	4	3	4	3	8	7	5	3	5
7	8	5	4	4	5	2	4	3	4
2	4	6	4	6	5	4	5	8	4
5	5	6	5	8	7	5	4	4	5
3	6	6	5	6	5	6	4	6	3
9	5	4	4	6	5	4	5	7	10
9	4	5	5	6	5	6	6	6	5
4	5	5	5	3	5	5	3	5	3

Mean = 5, RSD = 35%

Want to be Within 20%

7	6	5	6	1	6	4	5	4	4
5	3	8	2	3	4	6	8	6	4
6	4	3	4	3	8	7	5	3	5
7	8	5	4	4	5	2	4	3	4
2	4	6	4	6	5	4	5	8	4
5	5	6	5	8	7	5	4	4	5
3	6	6	5	6	5	6	4	6	3
9	5	4	4	6	5	4	5	7	10
9	4	5	5	6	5	6	6	6	5
4	5	5	5	3	5	5	3	5	3

Mean = 5, RSD = 35%

Want to be Within 20%

3	4	2	9	0	4	0	4	5	5
6	11	1	5	0	8	7	6	8	1
12	0	9	4	3	0	7	9	7	3
6	18	12	13	3	4	10	17	0	0
6	16	9	6	4	10	0	7	0	4
10	5	9	11	6	3	5	6	4	0
11	12	6	0	13	6	9	6	12	5
4	8	13	1	10	1	2	7	0	6
6	1	7	0	0	6	0	8	3	0
2	1	11	13	6	5	9	3	0	11

Mean = 5, RSD = 100%

Limit = 3

3	4	2	9	0	4	0	4	5	5
6	11	1	5	0	8	7	6	8	1
12	0	9	4	3	0	7	9	7	3
6	18	12	13	3	4	10	17	0	0
6	16	9	6	4	10	0	7	0	4
10	5	9	11	6	3	5	6	4	0
11	12	6	0	13	6	9	6	12	5
4	8	13	1	10	1	2	7	0	6
6	1	7	0	0	6	0	8	3	0
2	1	11	13	6	5	9	3	0	11

Mean = 5, RSD = 100%

Limit = 7

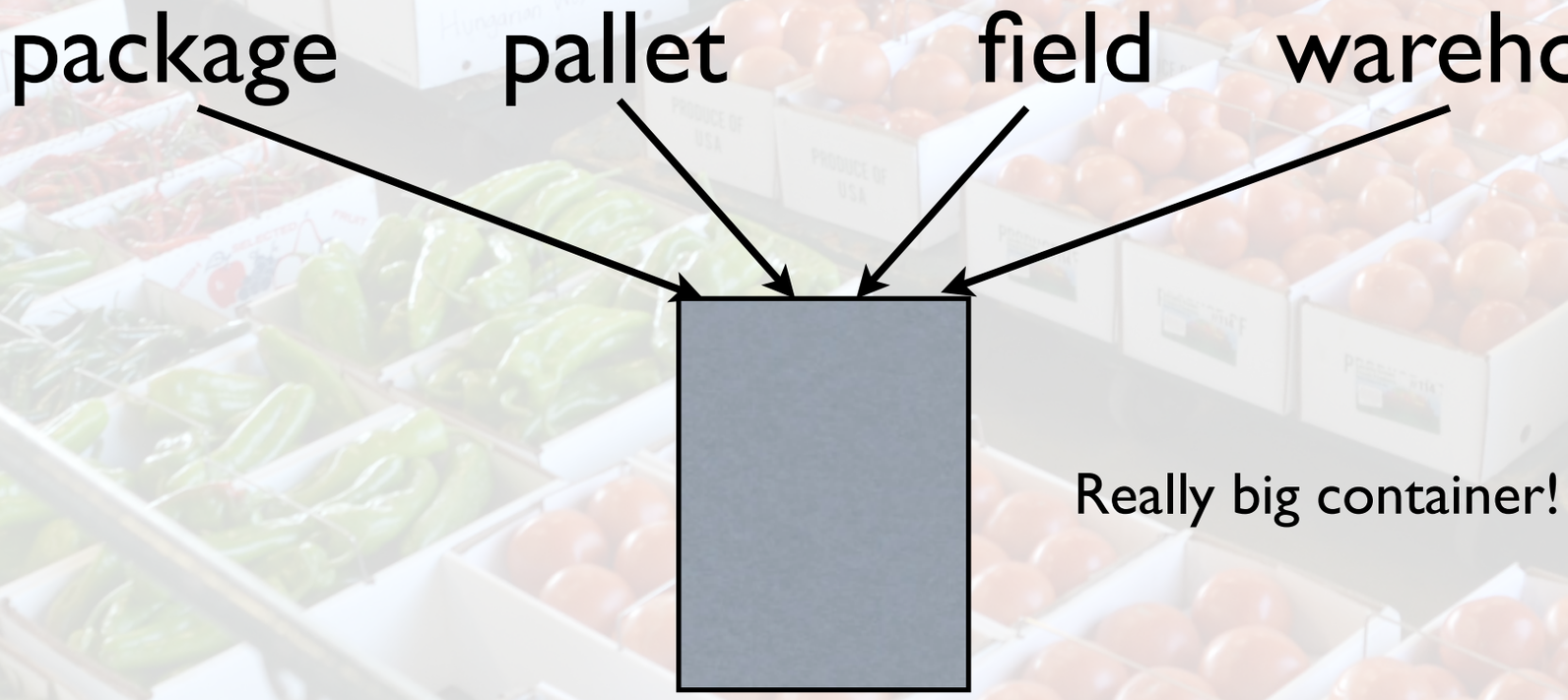
3	4	2	9	0	4	0	4	5	5
6	11	1	5	0	8	7	6	8	1
12	0	9	4	3	0	7	9	7	3
6	18	12	13	3	4	10	17	0	0
6	16	9	6	4	10	0	7	0	4
10	5	9	11	6	3	5	6	4	0
11	12	6	0	13	6	9	6	12	5
4	8	13	1	10	1	2	7	0	6
6	1	7	0	0	6	0	8	3	0
2	1	11	13	6	5	9	3	0	11

Mean = 5, RSD = 100%

Consequences of Error

- Incorrect estimation of true concentration
- Conclude it is below the specification limit when it is above
- Conclude it is above the specification limit when it is below
- Fail to detect presence of a banned substance
- Miss a problem (especially a rare one)

Ideal Sampling



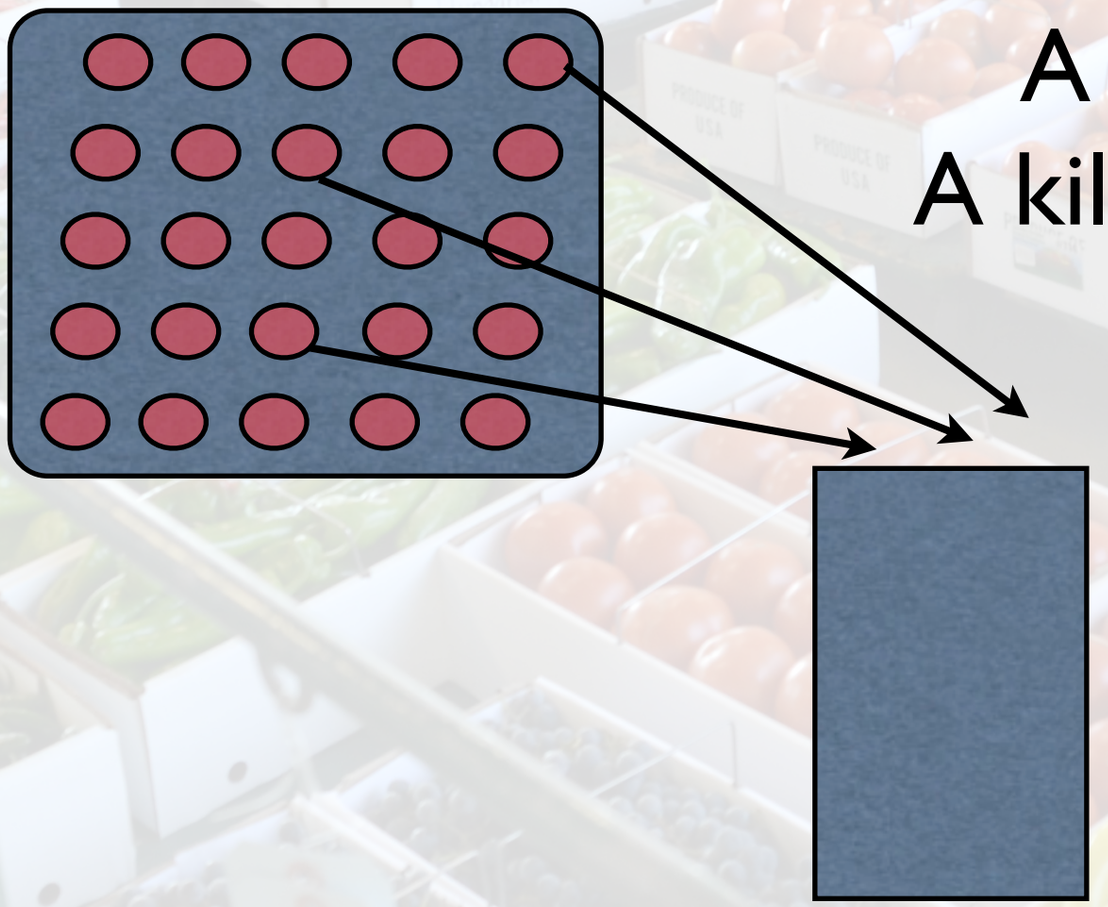
Really big container!

The perfect sample is to take it all!

What would the sampling error be?

Next Best Sample

(lots of material from lots of places)



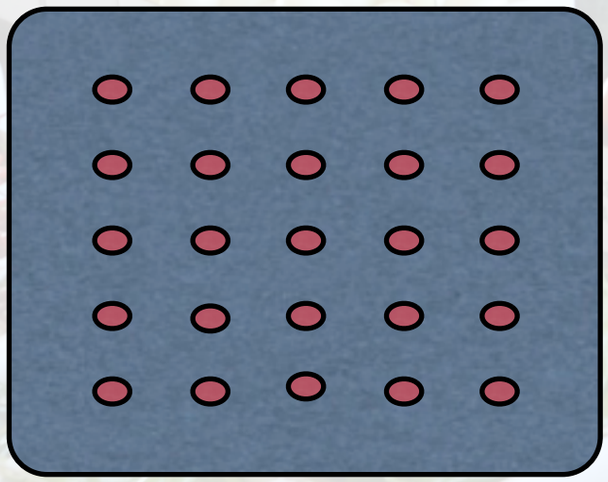
A liter every meter

A kilogram every meter

Really big sample jar!

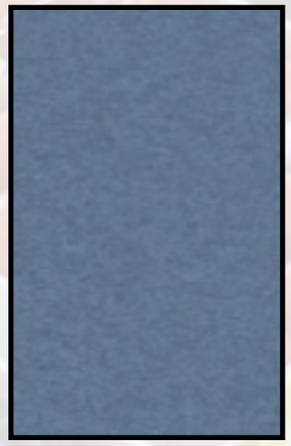
Next Best Sample

(a little material from lots of places)



100 ml every meter

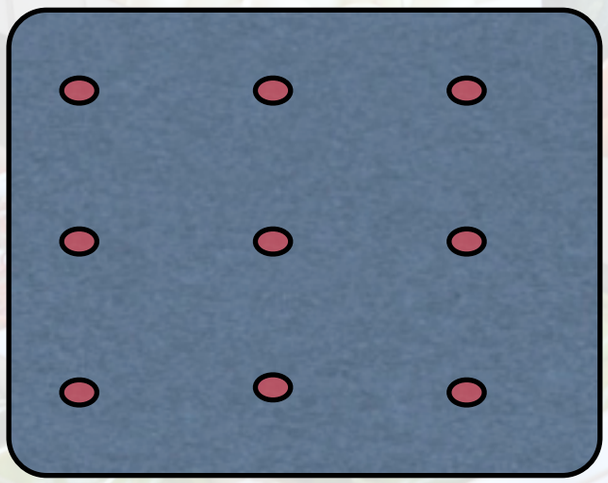
100 grams every meter



Not quite a really big sample jar!

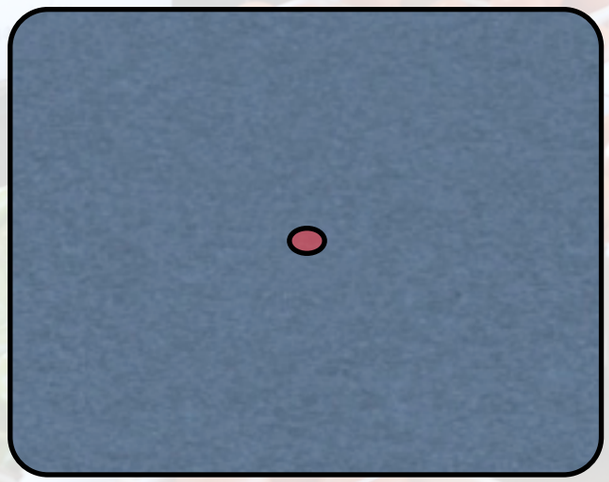
Next Best Sample

(a little material from a few places)



Regular sized sample jar

Worst Possible Sample

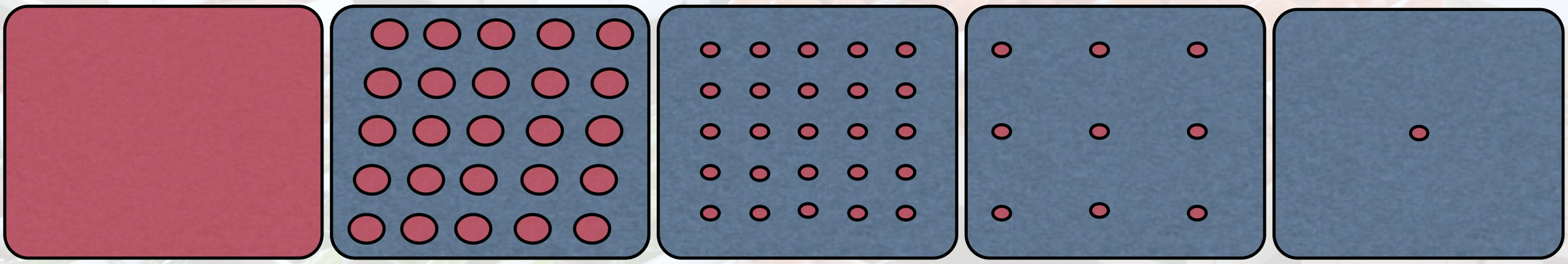


5.0 ml or
1.0 gram

Just collect a small amount from a single spot
AKA grab or discrete sample

This is true in the field as well as the laboratory

Options



What is sufficient for your project objectives?

Sampling Logic

- What is the material in question (decision unit)?
- Does the decision unit fit in the sample jar?
- If so, collect it all
- If not, collect a sample to get the same answer you would get had you put the entire decision unit in the jar
  - enough mass
  - enough increments
  - sample correctness (tools)
- If we cannot collect the entire decision unit, we will have sampling error (the ability for the sample to truly represent the decision unit)

Where to Begin?

- There are only two types of samples
  - representative (can make defensible decision)
  - non representative (cannot make defensible decision)
    - fit for purpose/not fit for purpose
- Need criteria to know if samples are representative/fit for purpose (SQC)
  - need to know question trying to answer
  - need to know what material is in question
  - need to know error rate (confidence)

Sample Quality Criteria (SQC)

- Number one reason incorrect decisions
- Needs to be done--and done correctly
- Should not be seen as a hardship, but as opportunity
- The key elements for sample protocol design
  - what is the question
  - identification of population(s) (decision unit)
  - confidence (how good does the decision need to be?)

Selection of the population seems to be the most problematic element for those new to the process

Question

- What is the analyte?
- What is the concentration of concern?
- Affects
  - types of tools and containers (esp trace levels)
    - losses
    - contamination
    - chemical change
  - sampling techniques
    - aseptic
    - oxidation
    - “clean” techniques
  - sample mass/volume

Question (cont)

- preservation
  - temperature
  - chemical
- holding times
  - shipping
  - time of collection
- health and safety
  - PPE
  - exposure time
- live organisms
  - replicate
  - die

Decision Unit

- Where increments are collected from
- Where inferences are made to
- Not too big, not too small, just right!

Decision Units

- How many are there?
- Can every one be sampled?
- If not, will a subset meet the objectives?
- Each decision unit is still sampled properly!
- Which do you regulate?
  - Is the lot of tomatoes within specs?
  - Is this package of tomatoes within specs?
  - Is this pallet of tomatoes within specs?
  - Is this tomato within specs?
  - Are all the fields of tomatoes in this county within specs?

Confidence

- How is the decision going to be made?
  - How will the data be used to make decision?
  - Are statistics going to be calculated?
    - need to be sure the right data is collected
- Is the error in the final result (confidence) going to be calculated?
  - If so, how? If not, why?
  - can't just include the analytical error
  - quality control

When you make a decision, do you want to be correct?

Bulk Sampling is Different

- Not the type of sampling encountered in
  - surveys
  - quality control
  - many food and feed documents (attribute)
- Random location is not sufficient to eliminate bias
- Number of samples is not sufficient to control error
- There is no fixed unit in bulk sampling
- Bulk sampling is more complicated

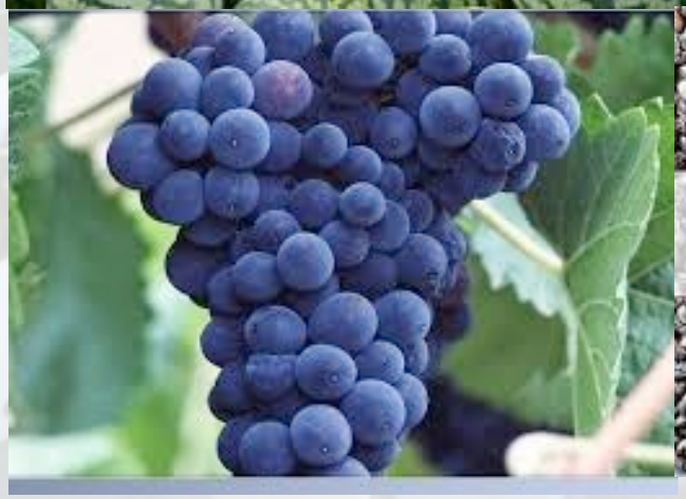
What is Bulk Sampling?

The taking of samples in arbitrary, irregular units rather than discrete units of uniform size for chemical analysis.

Free Dictionary by Farlex

When the individual fragments (molecules) are too small and too numerous to count. Therefore, it is impossible to individually identify and collect them at random. Sampling must be performed by collecting multiple fragments at one time.

No Choice as Elements Become Smaller



Attribute vs Bulk

- No fixed "natural" element in bulk sampling
  - water
  - mashed potatoes
- Could have fixed element in the field (and can take it all), but not in the lab
- May want to perform attribute testing (percent of compliant decision units) in the field, but cannot take the entire decision unit so need to bulk sample

What Information is Desired?

- Average of everything
  - What is everything?
  - How is everything defined? Decision Unit
  - How close to the true average do you want to be?
- Percent of individual items above or below
  - Which items does the percent apply to? Decision Unit
  - How close to the true percentage do you want to be?

Attribute sampling consumes a lot of resources

Controlling Field Sampling Error

- Need enough sample mass
- Need enough increments
- Need random selection of material
  - access
  - tools
- Need to maintain analyte integrity
  - sampling techniques
  - containers

Error is not controlled through the number of samples, but through quality of samples

Controlling Processing/ Handling Error

- Maintaining analyte integrity
  - preservation
  - transportation
  - holding times
- Maintaining sample representativeness
  - comminution
  - sieving
  - sample mass reduction

Analytical Error

- Can be large if poor laboratory practices
  - incomplete extraction
  - poor calibration
  - host of other things that can go wrong
- But we have...
  - systems and procedures
  - quality control
  - accreditation
  - training

Just like in the field. Right?!

Errors

- What is the big deal with errors? Does it really matter?
- What errors do we need to measure?
- How do we currently measure error?
- If we don't measure error, how can we certify "true and accurate?"
- How do we know how much error we can tolerate in our measurement system?

Errors Change with Concentration

- May be easy to find aflatoxin when there is high concentration, but can you find aflatoxin when there is only a minute quantity?
- Protocols need to be able to detect
  - when material is just fine
  - when material has upset condition
- Which is more important?
- Which condition does your protocol address?

Errors Change with Concentration

- The smaller the concentration (or percent defect), the harder the sampling problem (the larger the error)
  - presence of any GMO - What does that mean?
  - GMO at 0.9%
  - GMO at 20%
  - GMO concentrations in the ppm and ppb range
- Percent level concentrations are typically less problematic. Were current protocols developed around these levels?

Not all Analytes Behave Equally

- Different concentrations
- Different heterogeneity
  - some segregate
  - some clump
  - some change when handling material
- The more isolated the analyte, the harder it is to sample
- In cases where multiple analytes from same field sample, need to represent most difficult

Revisit a Few Topics

- Decision units
- Theory of Sampling
  - mass
  - increments
  - tools
  - integrity
- Quality control

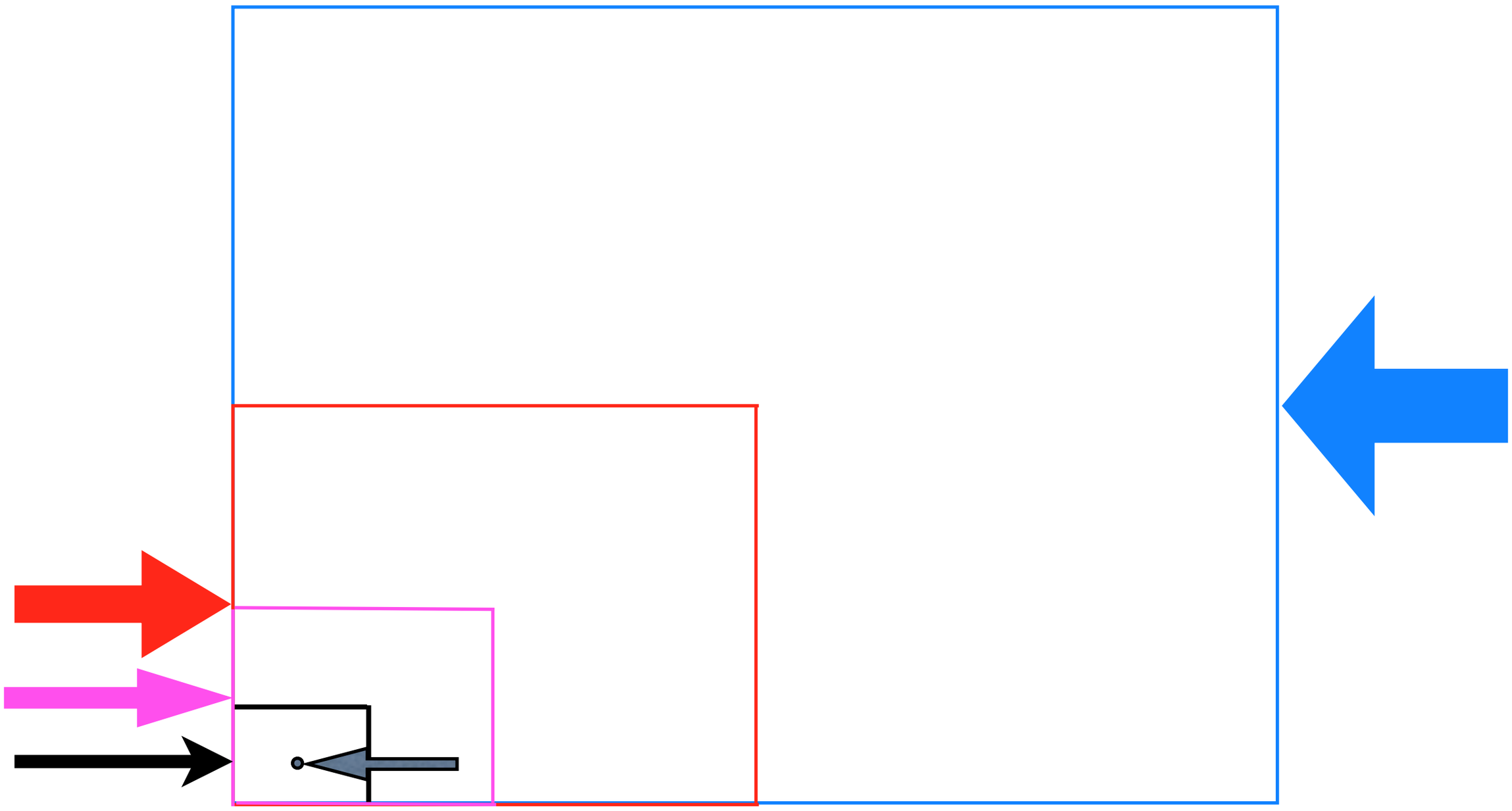
Decision Units

- For the average of "everything," the decision unit is "everything"
- For the percent of individual items, the decision unit is the item and there are many, many decision units
- Completely different information
- Completely different field sampling
- Completely different laboratory processing

Is Compliance Tied to Decision Unit?

- Label claim by manufacturer
- Regulatory limit
- Dose based criteria
- ??

May all be different decision units



What needs to be in compliance?

What is harder to accomplish?

What is Important for Dose?

$$DI = C \times IR$$

DI - daily intake (dose)
C - concentration (mass/mass)
IR - feed intake rate (mass/day)

Assuming an absorption efficiency of one, which may not be accurate

We must have the average concentration (C) of the intake amount for this formula to work

"All things are poisons, for there is nothing without poisonous qualities. It is only the dose which makes a thing poison." — Paracelsus

Is This a Problem?
Limit = 400 mg/kg

800 mg/kg

50 mg/kg

25 mg/kg

It is not about concentration, it is about dose!

Is This a Problem?
Limit = 400 mg/kg

292 mg/kg

This is just the average of the three smaller spoons.

Is This a Problem?
Limit = 400 mg/kg

200 mg/kg

50 10 550 10 480 100

800 mg/kg

100 1600 50 1500 70 1480

If took a small discrete sample (i.e., a gram in the laboratory), what is your error rate?

When Making Decisions...

- The scale of observation is as important as the analytical result!
- Knowledge of the error is as important as the analytical result!
- Data is meaningless without the objectives as a framework to evaluate

Manufacturing

- What does facility certify?
- Is that consistent with their sampling?
- Would I even bring it up if it was true?!

Sampling Theory

- Mass (particle size and nuggets)
- Increments (random)
- Tools (shape and extraction)
- Integrity (sample and analyte)

Orange	Yellow	Red	Brown	Green	Blue
20	14	13	13	16	24

What is the minimum number of M&Ms we need?

What about a trace amount (0.1%) of AFDO M&Ms?

Should M&Ms be sampled individually or by the handfuls?

Mass and Error

- Take multiple 100 gram samples
- Take multiple one pound samples
- Take multiple ten pound samples
- What will be the result?
  - ten pound - lowest RSD
  - 100 gram - highest RSD
- Why?
- Consequences?

Out Of Spec and Decision Units

- The OOS criteria must be based on the "compliance" unit
- Sampling a unit too small will result in too many false OOS
- Sampling a unit too large will result in fewer OOS than there should be

Increments at Random

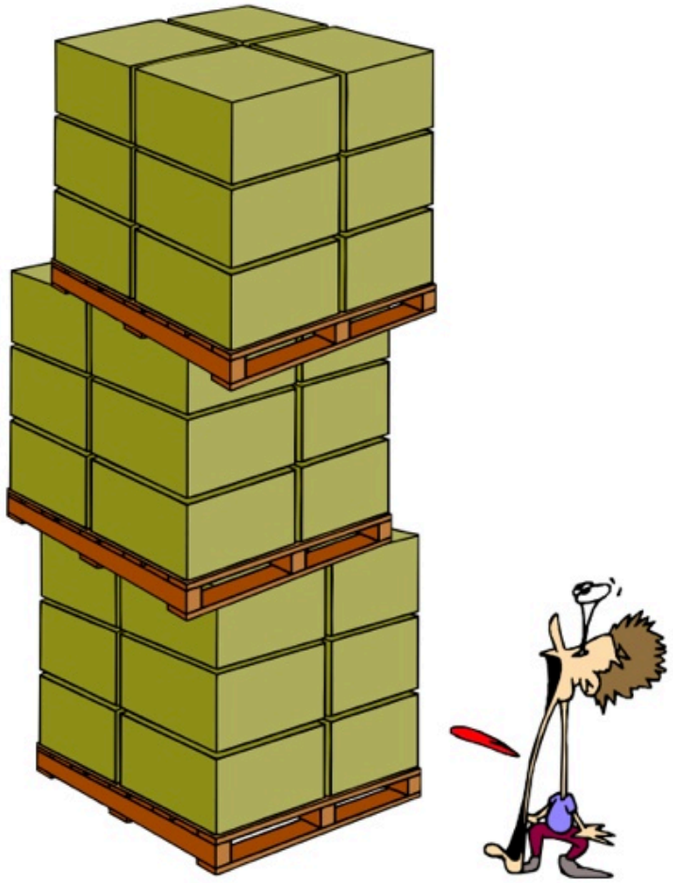
- Or no ability to make inference
  - simple
  - stratified
  - systematic
  - host of other
- Some random schemes are easier to implement than others
- If material is selected because it is different
  - What can be stated about the Decision Unit?
  - Does that meet the SQC?

Inference

- Only to the part that was available
- Only to the random portion of available
- Only what was in the container if biased
- Can use professional knowledge
  - but what is it based on?
  - not a good habit to get into
  - not equivalent

Very Important

Accessibility



"If the nature of the presentation of the lot…makes it impossible to adequately apply these procedures…an alternative presentation of the lot [should] be sought"

- International Rules for Seed Testing

Doug Miller, RGT
Illinois Crop Improvement Association
217-359-4053 dmiller@ilcrop.com

Tools

- Design criteria
  - parallel sides for stream
  - three times opening
  - collection of all particles intended
- Observation and thought go a long way
- Tools make the decision unit more accessible
- Some of everything in the correct portion
- Equiprobable selection

Integrity

- What is the concentration in the Decision Unit?
  - sampling did not change it
  - packing and transportation did not change it
  - laboratory processing did not change it
- Multiple analytes (behave differently)
- Many factors
  - environmental factors
    - humidity
    - temperature
  - time
    - holding time
    - preservatives

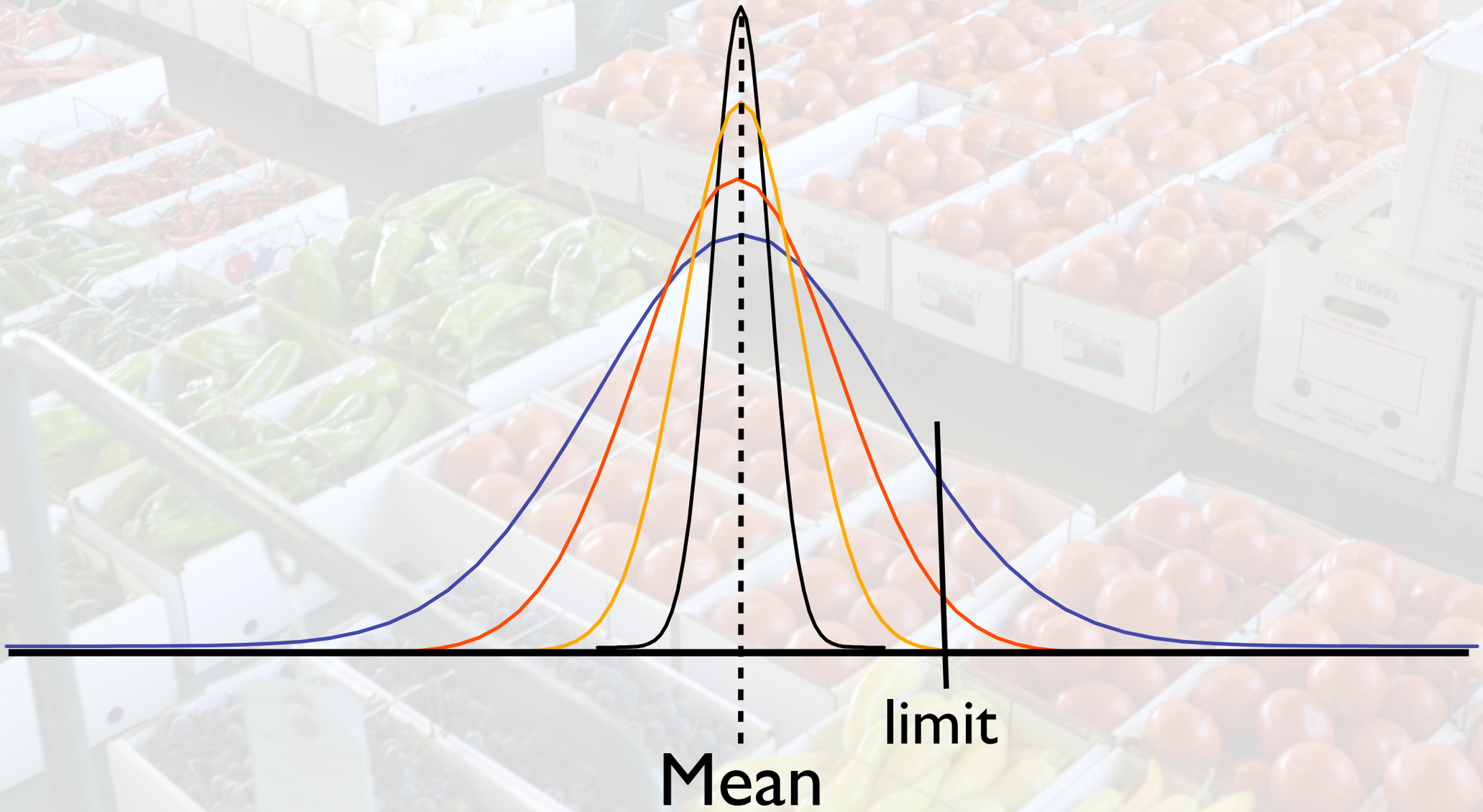
Don't Forget Quality Control

- Consequences of decision
- Distance from regulatory threshold
- Accreditation or other program requirements
- Knowledge of error important
- Depends on claim to be made (SQC)

Error in the Entire Measurement System

- Many data users want to know the total error in the results presented
- When error calculations are presented many data users think it refers to the total error in the result
- Does it?
- Does the decision maker need to know the error in the entire measurement system?

Consequences of Not Controlling Error



Quality Control

- What quality control is performed in the laboratory?
- Why is quality control performed in the laboratory?
- What would be the consequences of not performing this quality control?
- Should all steps in the measurement process have quality control or only those with the largest potential source of error?

Lab vs Field QC

- More relevant in the field than in the lab
  - less controlled environment
  - larger potential source of error
- Can be harder to implement in the field
- More critical with new and emerging contaminants
- Must at least be considered in any sampling campaign

Quality Control

- Necessary to determine
  - if sampling is in control
  - the magnitude of the error
- If it is important in the lab, it is even more important in the field
- What QC is currently performed in the field?

answer:

Quality Control

- It involves effort in the field
- How often does it need to be done?
  - every time
  - certain percentage of the time (e.g., 10%)
  - enough to prove a method
  - critical samples
- A balance between risk you are willing to take and resources you are willing to spend
- Philosophically, how can you spend tremendous resources in the lab that produces only a small portion of the error and no resources in the field where the errors are large?

Types of Quality Control

- Contamination checks
- Precision (error) measurements
- Post sampling comparisons (splits)
- Nothing for bias in sampling
- No sampling reference materials

What Do We Need to Improve the Quality of Data?

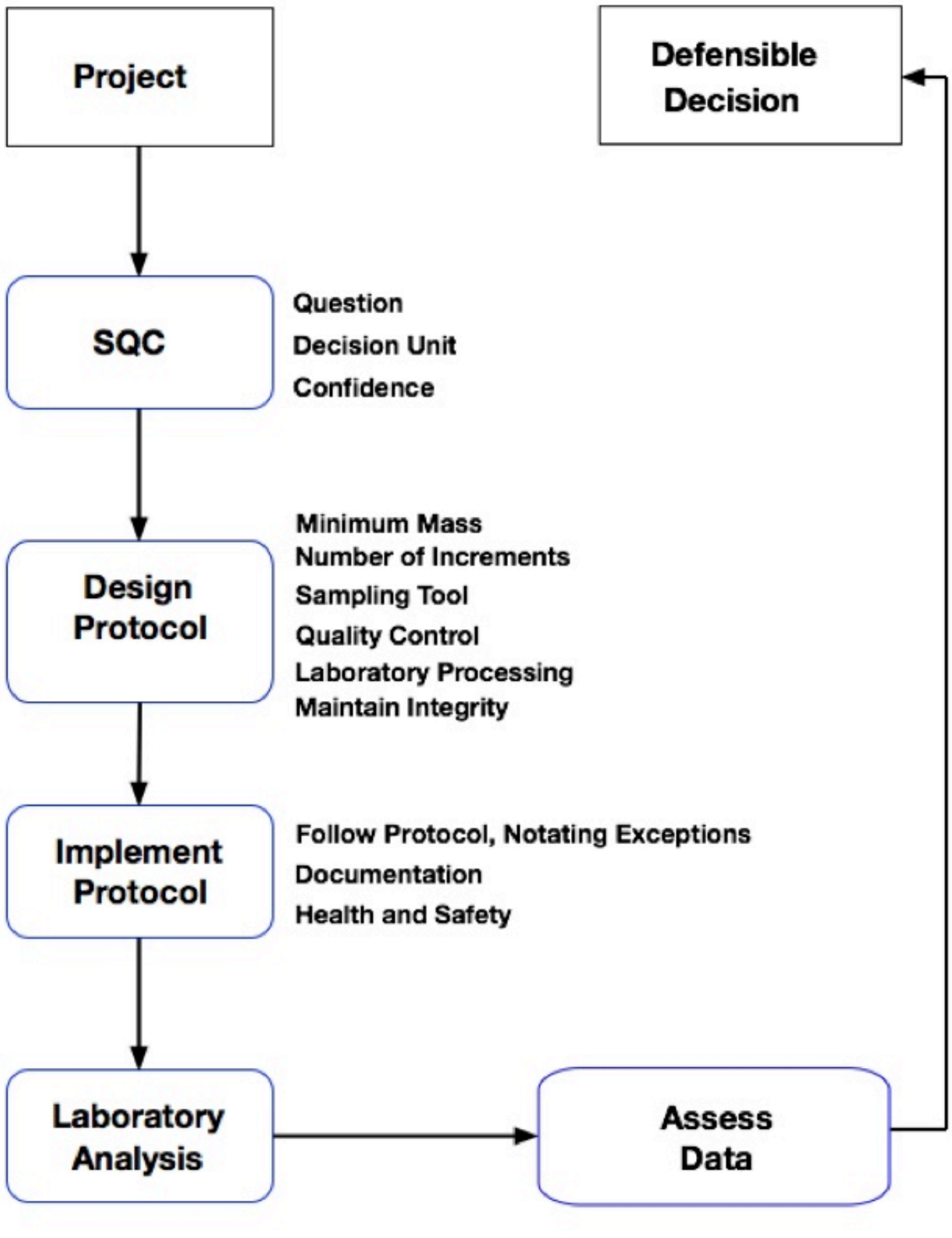
- Systematic approach for all sampling
  - equivalency of data
  - defensible decisions
- Power to make good decisions
  - if we did not find it, was it there?
- Communication
- AND....
- ...GOODSamples

What Needs to be Done?

- Put systems/procedures in place to define, detect, measure and control **all** errors
- Will these sampling procedures be identical to analytical procedures?
- Who is responsible for implementing these systems/procedures?
- **Who is responsible** to estimate the total error in the analytical result?

GOODSamples: Guidance on Obtaining Defensible Samples

Sampling and Sample Handling Working Group
FDA, AAFCO AFDO, APHL,
JANUARY 2014
web url

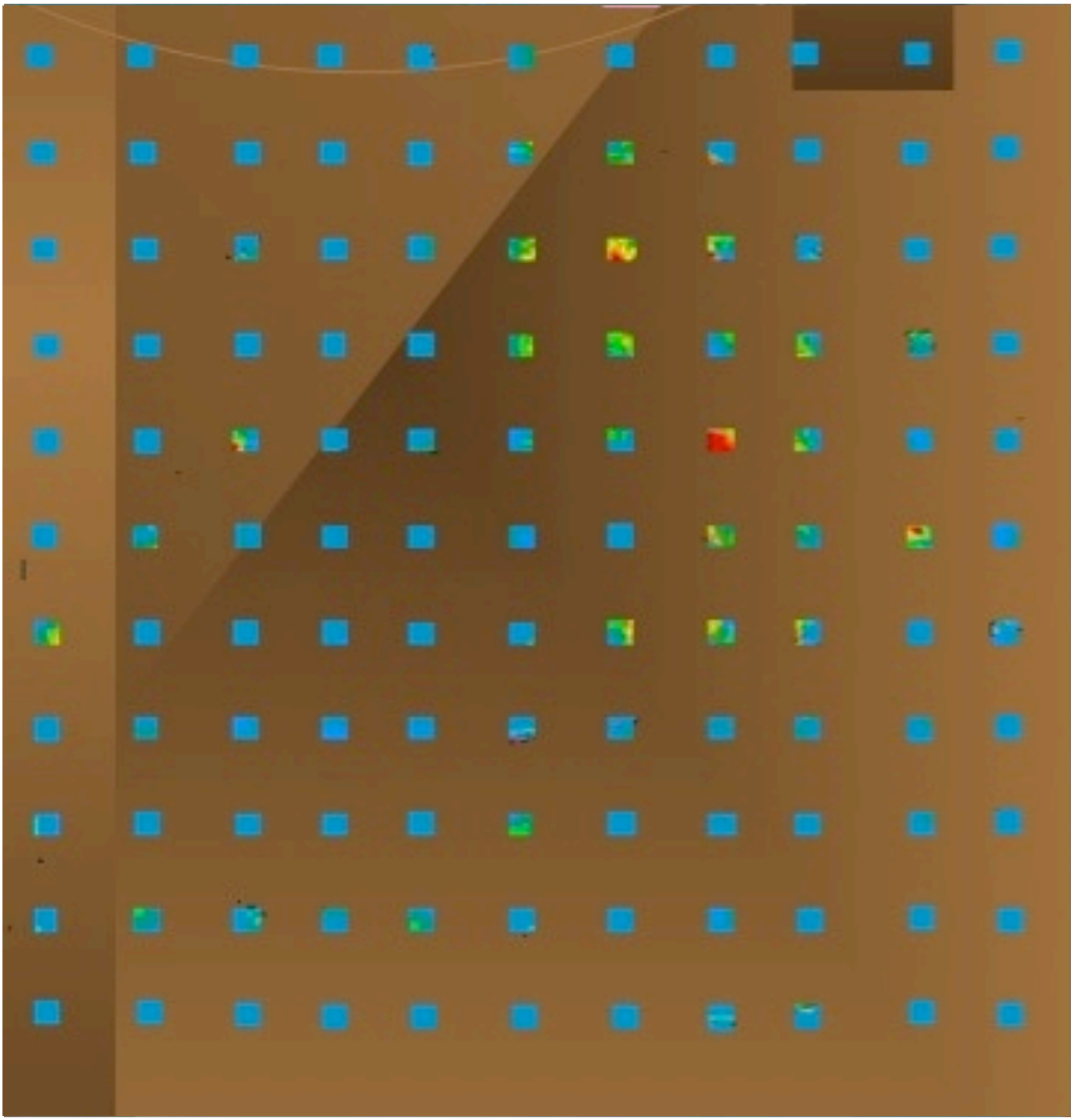


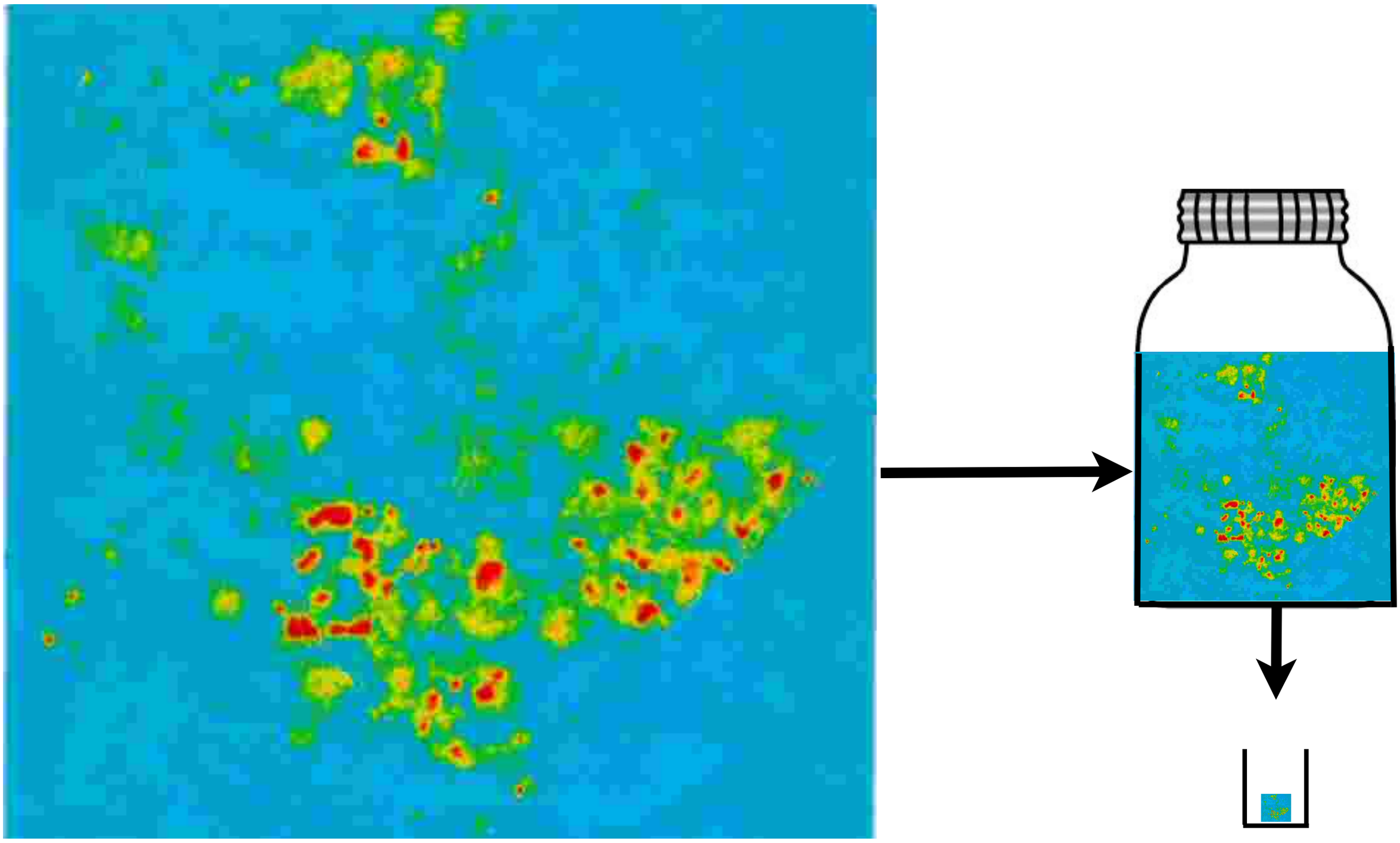
In the End There Are Two Types of Samples

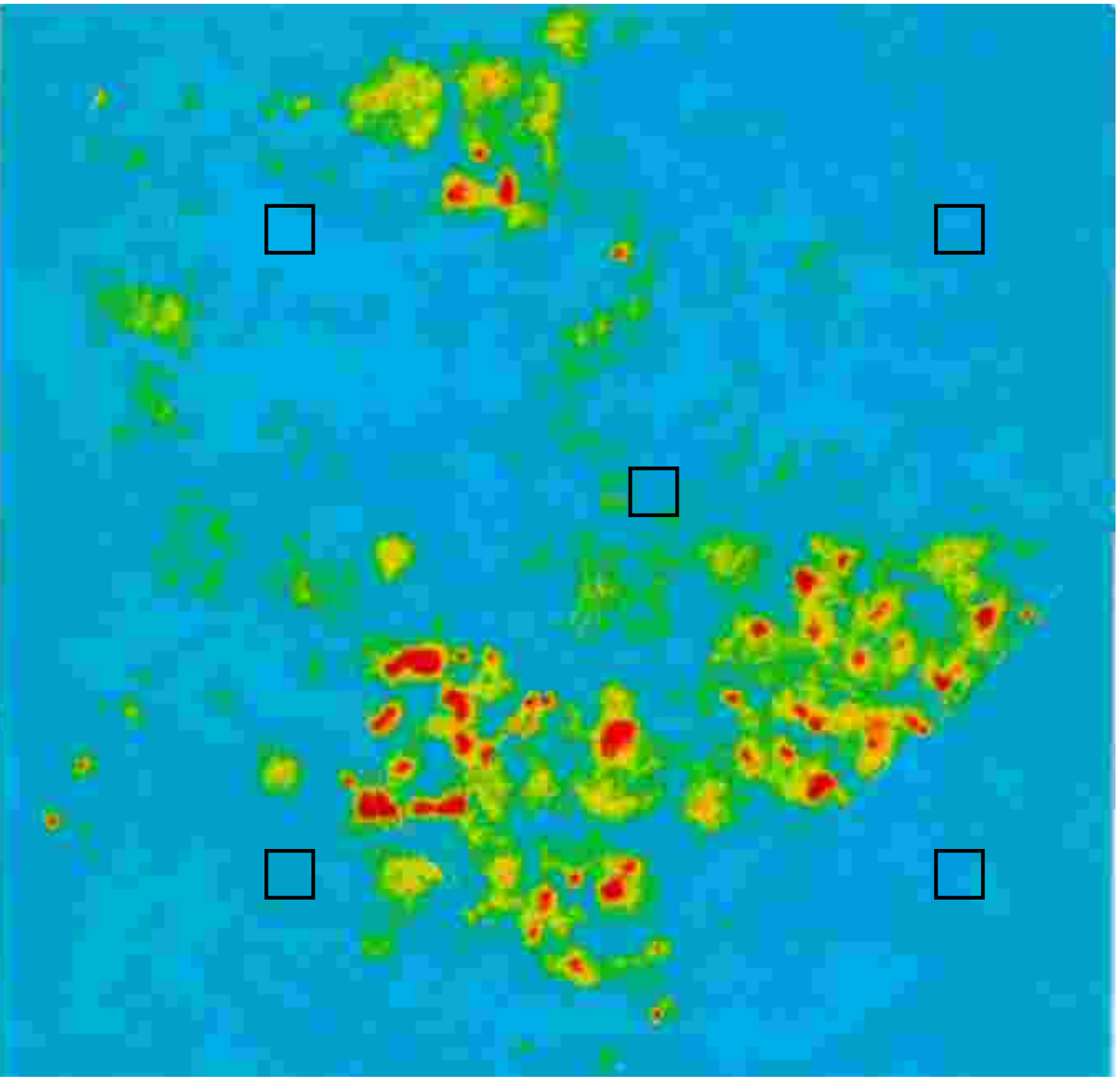
- Representative
- Non representative
  - convenience
  - judgment
  - single increment

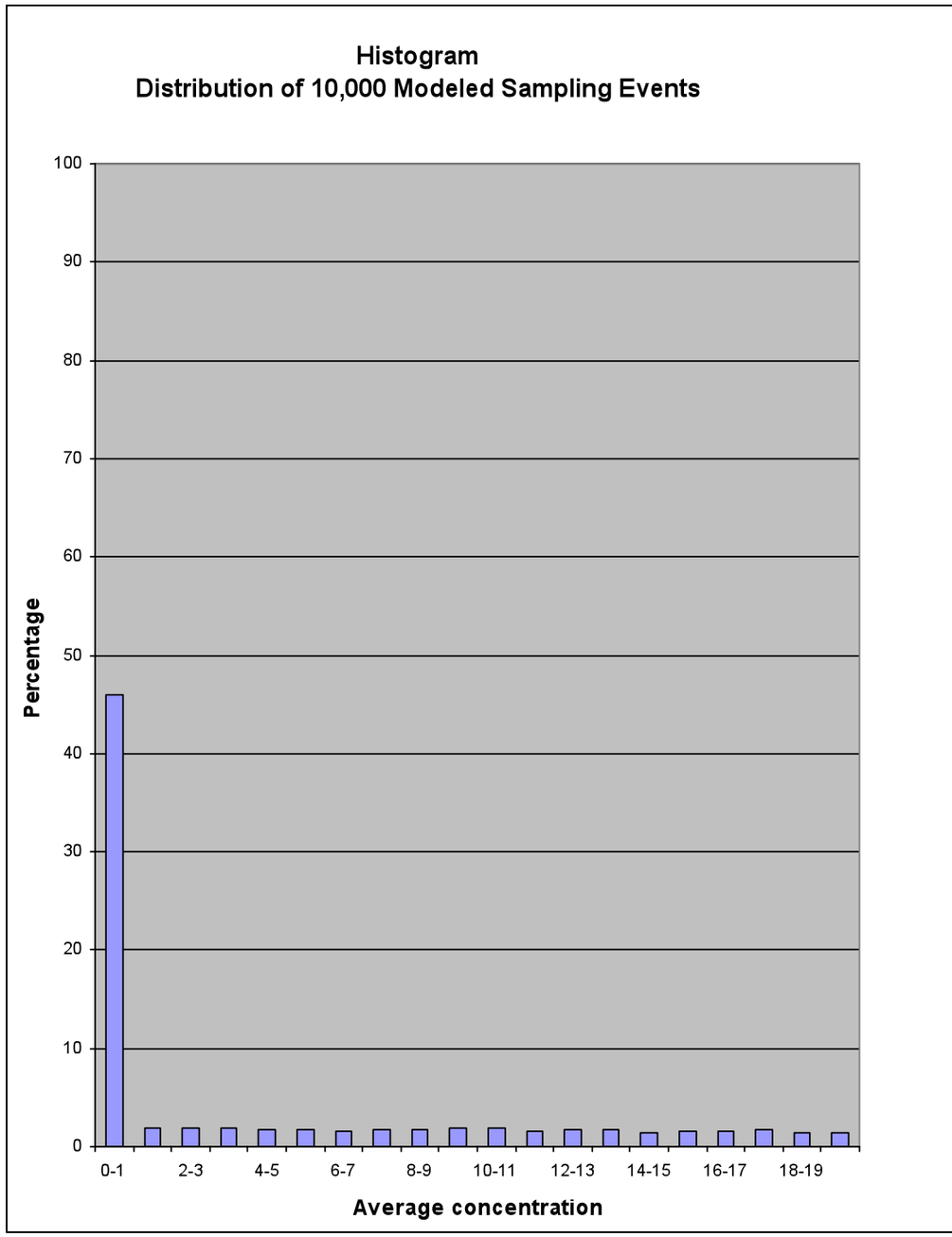
There may be many purposes, but only two types!

"An inaccurate sample ... may actually be worse than no sample at all." Steven Seideman









Miss almost 50% of time

Representative Sampling

- What is in the Decision Unit is in the primary sample
  - some of everything
  - same proportion
- The integrity of the analyte is not compromised
- What is primary sample is in the test portion
  - some of everything
  - same proportion
- We can demonstrate that it is so
  - documentation
  - quality control

as opposed to...

Kim H. Esbensen

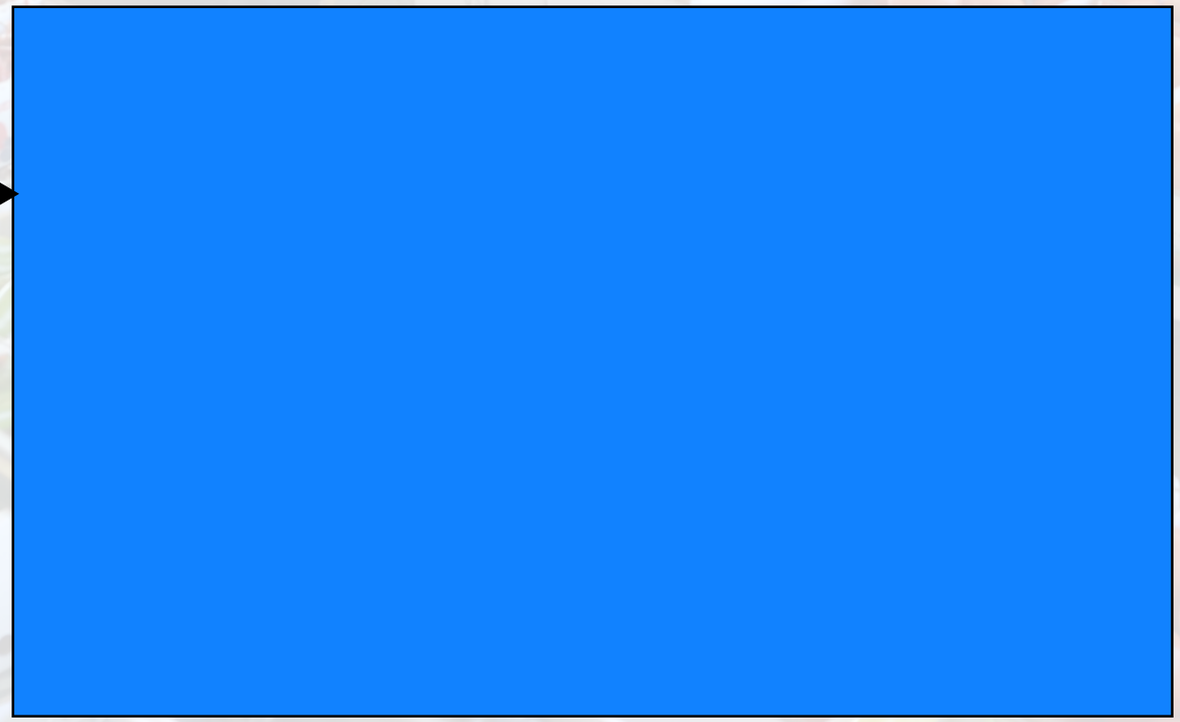
ACABS: Applied Chemometrics, Analytical Chemistry, Applied Biotechnology, Bioenergy & Sampling research group, University of Aalborg Esbjerg (AAUE)
Denmark

One Final Caution
Chasing the Molecule
Beware!

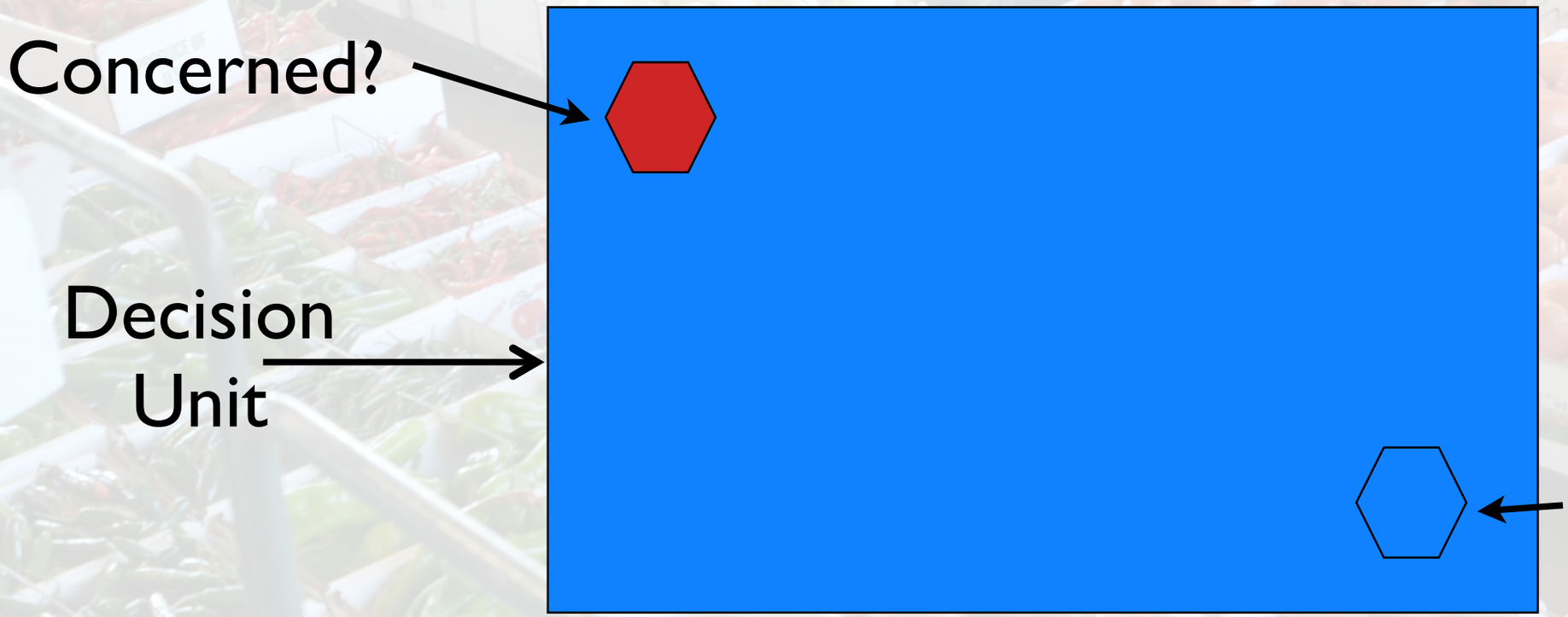
- Even if it exists, can you detect it?
- What is the meaning of one molecule?
- Need to have the Decision Unit "scale of decision making" to get defensible answer
- Don't miss the forest for the trees!

How Would You Sample This Decision Unit?

Decision Unit →

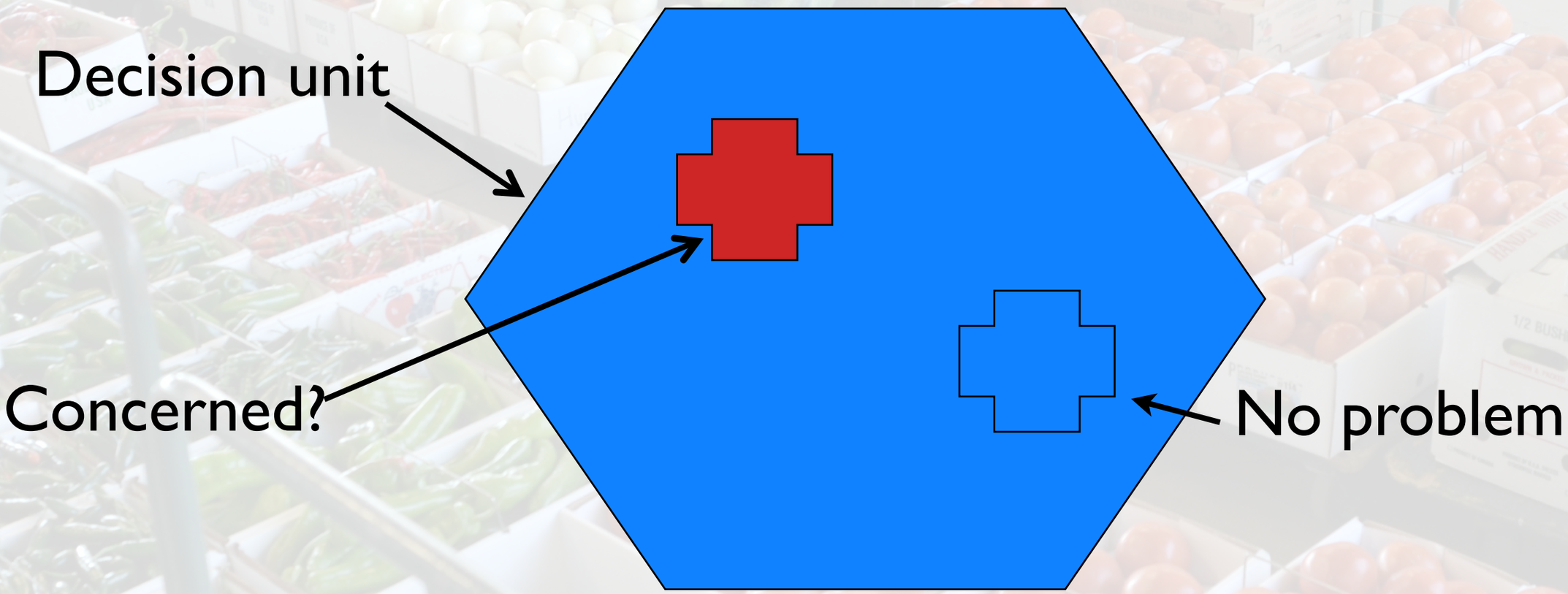


Is This a Problem?



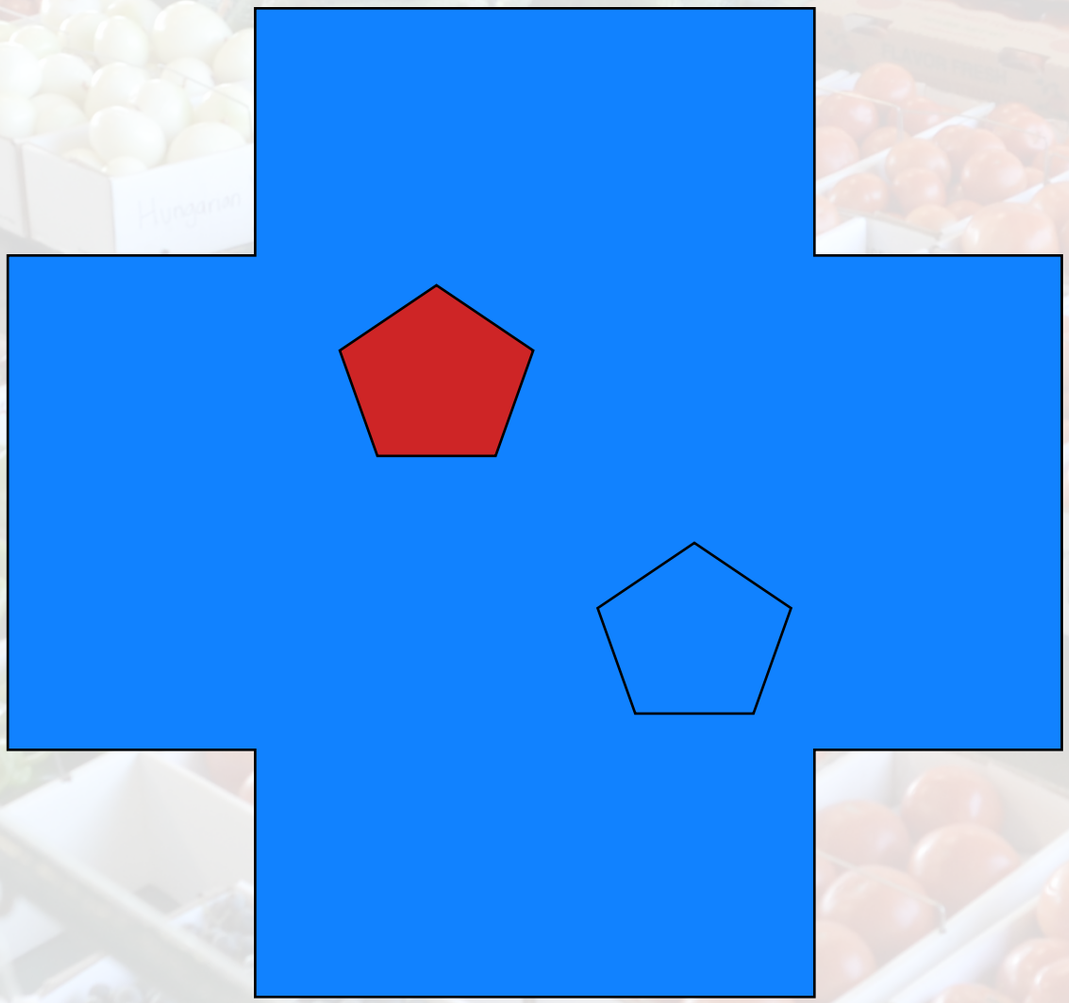
Area in red above the limit.
The average is within the limit.

Remember, No Problem



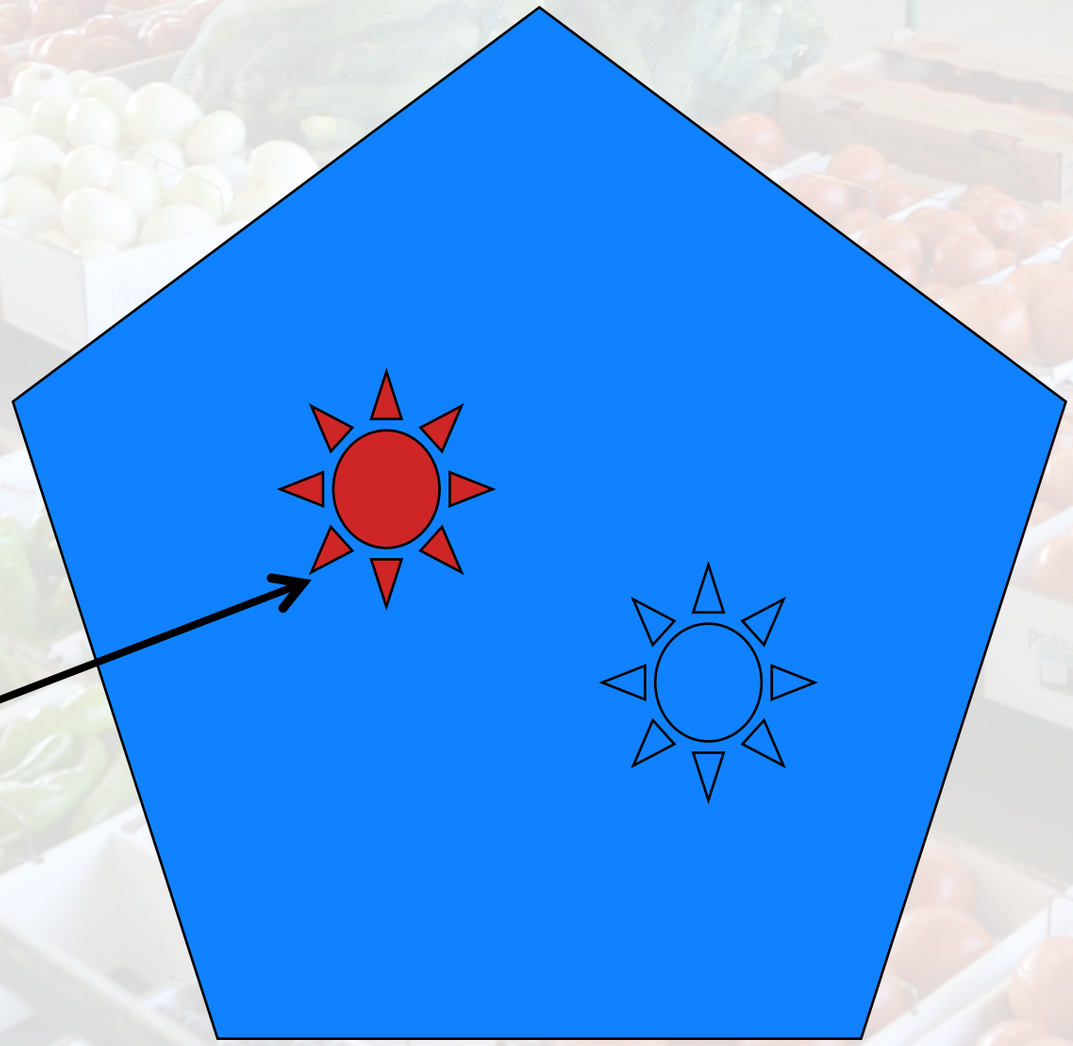
Area in red above the limit.
The average is within the limit.

Quit Changing Your Mind!

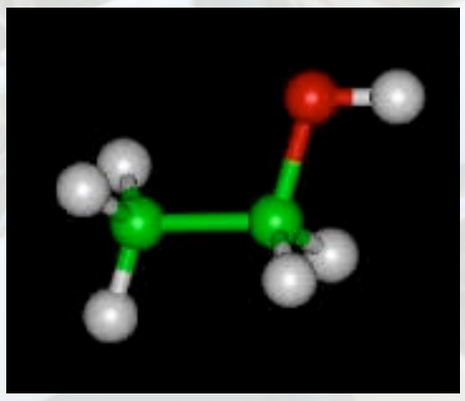


Area in red above the limit.
The average is within the limit.

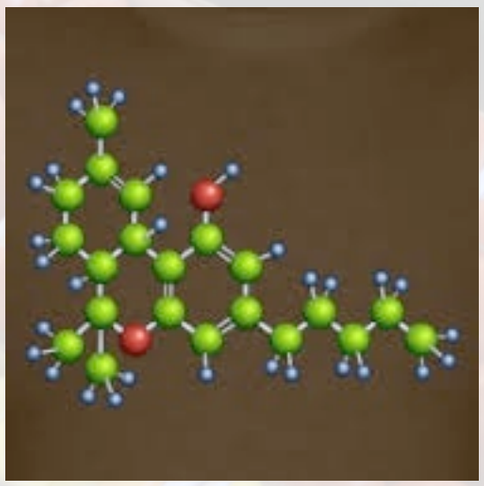
Finally, the Molecule!



Concentration
1,000,000ppm!



Area in red above the limit.
The average is within the limit.



Sampling Bulk Materials

- Must be tied to a decision unit (as with all sampling)
- No natural distribution (artifact of error)
- More complex than attribute sampling (where the number of random samples is all that matters)
- More opportunity to control sampling error and reduce the required sampling and analytical resources

Molecular/Atomic Scale

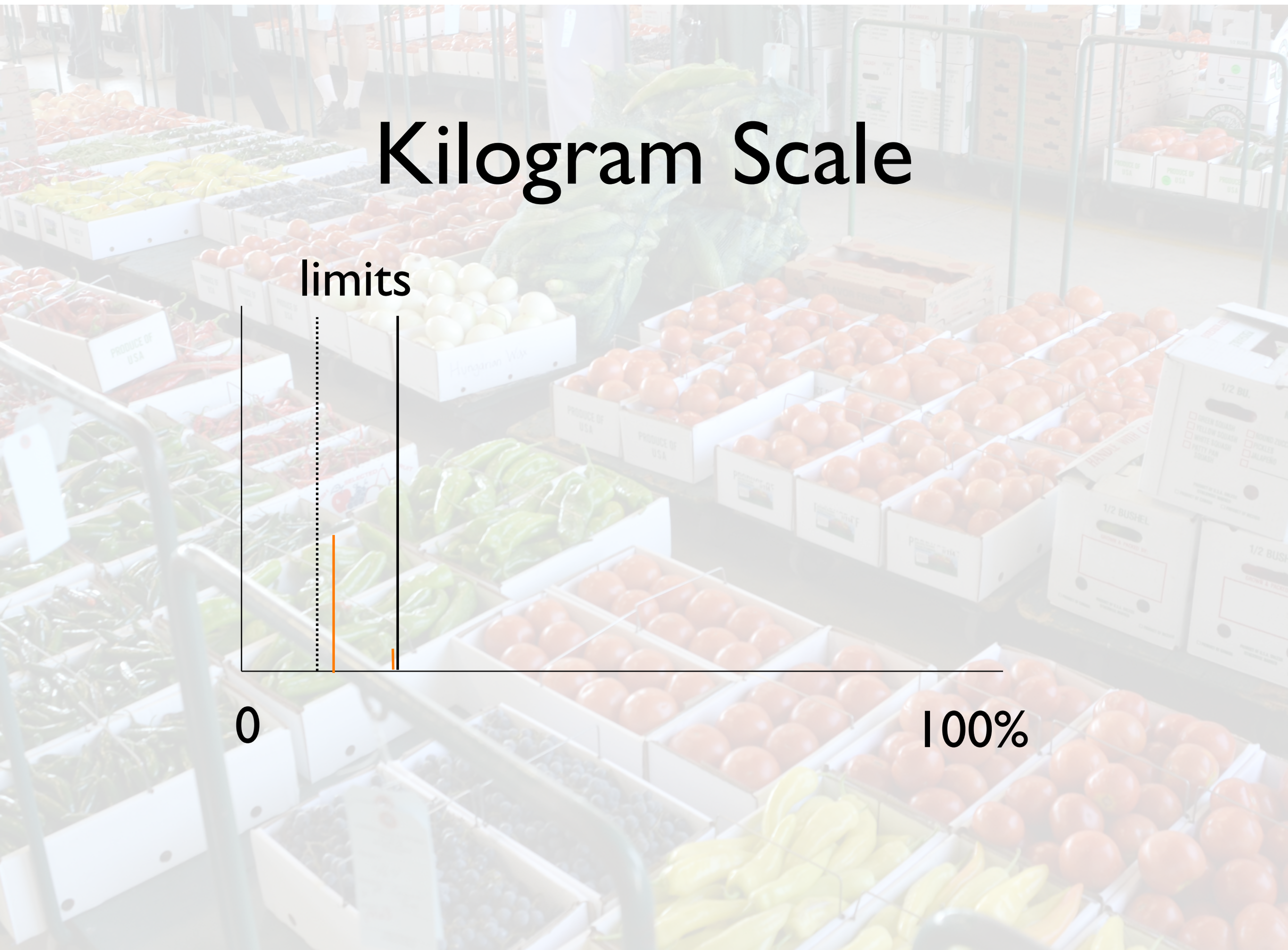
Smallest support possible

(Bimodal distribution)

Particle Scale

Gram Scale

Kilogram Scale



Entire Population

limits

0

100%

Largest support possible

Single value - truth

If I Were to Question Your Data

- Is your sampling adequate for your current objectives?
- What is the error in your final number?
- Where is your field quality control?
- Are you using proper equipment and is it in good working order?
- Is your sampling program resource optimal?

Resource Management

- Do all analyses cost the same amount of money and take the same time?
- Are all samples collected the same way?
- Some results have less consequences so can tolerate more error
- It would be a mistake to spend the same resources on easy and hard sampling problems

Consequences of Error Changes Closeness to Limit

- If the result is a factor of 10 different than the limit can a 25% error be tolerated?
- If the result is within 10% of the limit can a 25% error be tolerated.
- If there is no limit, how much error can be tolerated?

GOODSamples is the Beginning to

- Defensible decisions
- Resource optimization
- Equivalent data

GOODSamples

- Is not installed and then "check the box"
- Is a new way of life/approach to problem solving
- Is only the beginning
- Is about time!!

Without GOODSamples

- More of the same
- Is that satisfactory to meet the demands
  - FSMA
  - your agency
  - the public
  - accreditation
  - you

Equivalent Data

- Will the data match?
  - sample for a shift at the factory
  - sample a pallet at the warehouse
- Will the data match?
  - take some material out of every 5th bag for 100 bags
  - take one pound sample at the retail outlet
- Will the data match?
  - State A takes 10 increments out of easy bag
  - State B takes 10 increments out of random bags
- Can we make the data equivalent?
  - Yes, but under what condition?

Observations

- Mass
  - minimum mass for sampling error not considered
  - current mass guidelines may be acceptable for some analytes and products, but not all
- Increments
  - important in the field and in the laboratory
  - not always at random throughout the decision unit
- Tools
  - adequate tools are available, but...
  - need to ensure access to all of the decision unit

Observations

- Integrity
  - good for the most part
  - need to consider new analytes and trace levels
- Documentation
  - good for the most part
- Laboratory subsampling and processing
  - good for the most part
  - some splitting and sample mass issues
  - single increment

Observations

- Sample Quality Criteria needs work
- Question - analyte and level
  - good
- Decision Unit - population of interest
  - needs work, what really is in question, is that what is being sampled
- Confidence
  - needs work, need to know error in our sampling and more importantly, the error in our decisions

How Would You Respond to...

- bio terrorism
- nuclear fallout
- emerging analytes

What is the first thing to do?

Opportunity is Knocking

- Thanks to FSMA we have an opportunity to improve our sampling to obtain more defensible data and make better decisions.

We need to take advantage of this and incorporate

- Sample Quality Criteria
- sampling theory
- quality control
- assessment of all error in the measurement system

Science

Starts with the Sample.

ESS, the leader in PreCleaned Certified™ Sample Containers, and the pioneer in PrePreserved® Sample Containers, is the trusted source when science matters.

- ESS VOA vials feature a "low bleed" septa (lot tested by Headspace Method TO15).
- PreCleaned Certified™ and Quality Certified™ Containers are Prepared and certified to meet new ELAP requirements (DOD/DOE).
- products included, TOC Certified Vials, 53 "Clean Room" Containers, PrePreserved® Containers and