



Assume nothing? Is lack of evidence evidence of lack when interpreting staging data?

Capturing diagnosis and stage
UKACR / NCIN conference
16th June 2011

Sally Vernon

With thanks to many others at the WMCIU, especially Tim Evans and John Broggio

Deputy Director
West Midlands Cancer Intelligence Unit

Tel: 0121 415 8128 Fax: 0121 414 7712
E-mail: sally.vernon@wmciu.nhs.uk



Overview

- Why do we want better staging data?
- How could making assumptions help?
- When can we make assumptions?
- How much do we trust the assumptions we make?

Improving Staging Data

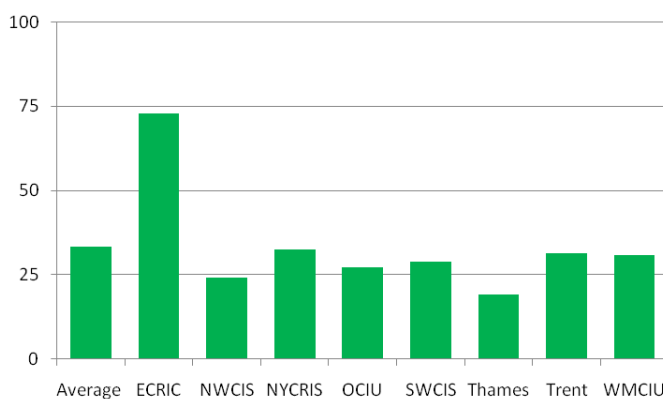


The completeness of staging data is an extremely important issue. I have been talking to the registries and agreeing with them that over the next one to two years we really do have to sort this issue.

– Professor Sir Mike Richards, to the Public Accounts Committee, December 2010

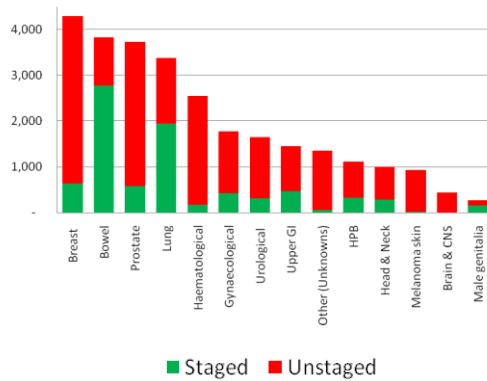
- Identifying late diagnosis
- Understanding treatment pathways
- Understanding case mix
- Measuring outcomes fairly

Current Staging Completeness



- UKACR 2011 Performance Indicators
- All malignancies excluding non-melanoma skin cancer
- % with known stage

Current Staging Completeness



- Stage completeness varies by cancer site in the West Midlands
- Only colorectal cancer appears to be well staged

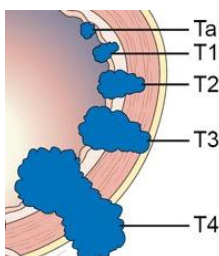
What is a stage?



TNM stage has three components

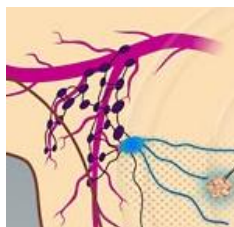
T

The **size** or extent of the primary tumour



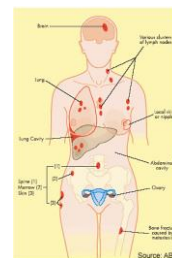
N

Absence or presence of malignancy in regional lymph **nodes**

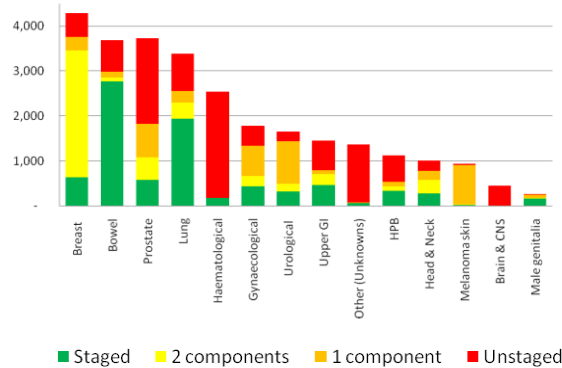


M

Absence or presence of distant **metastases**



Partial Staging Data



- Entirely incomplete staging data is uncommon
- Some cases have a T stage, or a T and N stage, but the rest of their staging data is missing

Can we make assumptions?



We know that the tumour is small, T1

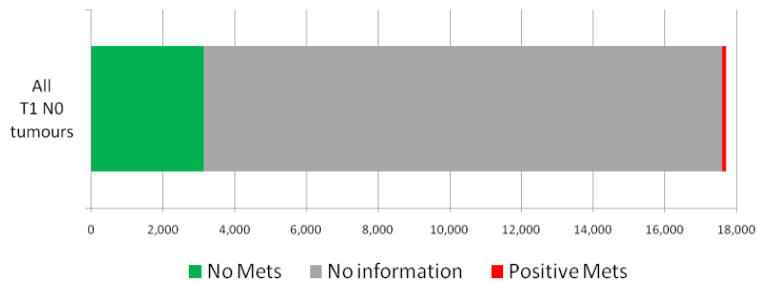
We know that the nodes were examined, and they were negative

We know that the tumour was excised, and the treatment looks curative in intent

We've got 50 pages of patient notes that don't mention metastases anywhere

So can we assume this tumour had not metastasised?

Even good assumptions fail



- Looked at all T1 N0 tumours over past 10 years
- Only 18% had data on whether metastases were positive or negative
- Metastases were negative 97% of the time where there was data
- But there were still 100 tumours with positive metastases
- Even for these very small tumours which have not spread to the nodes, the assumption fails sometimes.

Are blanks missing at random?



An example:

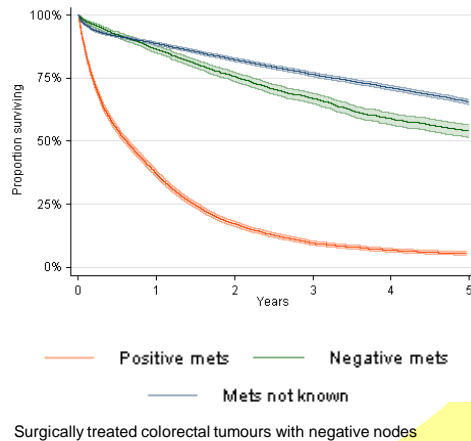
- Take a group of tumours
 - All surgically treated colorectal tumours with negative nodes ^[1]
- 1,109 where the metastases were stated as negative
- 482 where the metastases were stated as positive
- 11,000 tumours where we had no data on metastases
- If the blanks were missing at random, would expect 70% to be negative and 30% to be positive.
- But look at the survival curve...

^[1] Because it demonstrates the point clearly – not all groups of tumours show this!

Are blanks missing at random?



- Patients with negative metastases survive better than those with positive metastases, as expected
- But patients with *unknown* metastases survive better than both!
- This does not look like a mixed group of patients where 30% of the cases have metastases
- Suggests that lack of evidence is evidence of lack!



So is lack of evidence evidence of lack?



- Not always sensible to make assumptions but seems reasonable for some groups of tumours
- Might be happy to make assumption for T1 N0 surgically treated colorectal tumours... but not for T4 N1 prostate cancers with no recorded treatment!
- Need to find the cases where we can make assumptions based on the other data we know already

Identifying 'sensible' groups

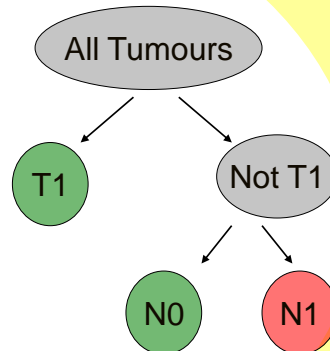


- Used statistical analysis
- Logistic regression can identify which variables are the best predictors of the presence or absence of metastases based on the known cases

Reassuringly, these turned out to be 'sensible sounding' things

like T stage, N stage, whether the tumour was surgically treated...

- Built a decision tree based on these predictor variables
- The end-points of the tree put tumours into the best groups for discussing whether or not we can make assumptions.

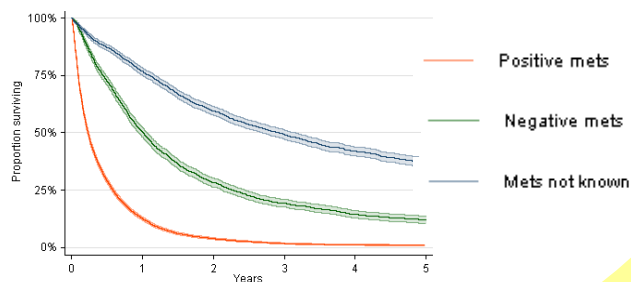


When would we assume negative metastases?



Four tests on each 'node' of the trees:

- 1) Look at the survival curves – are the unknowns like the known negatives (or even better!)?



Surgically treated lung cancers with histology of primary provided

When would we assume negative metastases?



Four tests:

2) Look at the output of statistical software which produced the tree

Probability (rule correct) = 91%

How happy would we be with this? Predicting 1/10 as having no mets when they had mets would affect data quality

But this prediction does not have extra information (eg the survival curves, the fact that data might not be missing at random)

So if the other tests are supportive, could accept any > 80% ?



Surgically treated lung cancers with histology of primary provided

When would we assume negative metastases?



Four tests:

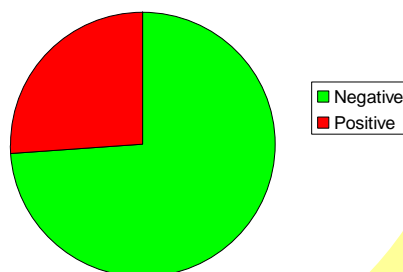
3) Look at the ratio of the known negatives to the known positives

Most of these tumours were negative

But 1 in 4 were positive

Don't want to make an assumption that is wrong 1 time in 4

But don't believe missing at random



Surgically treated lung cancers with histology of primary provided

When would we assume negative metastases?



Four tests:

4) Does it make common sense?

Surgeons would be unlikely to attempt a curative resection on a lung cancer which had metastasized at diagnosis



Surgically treated lung cancers with histology of primary provided

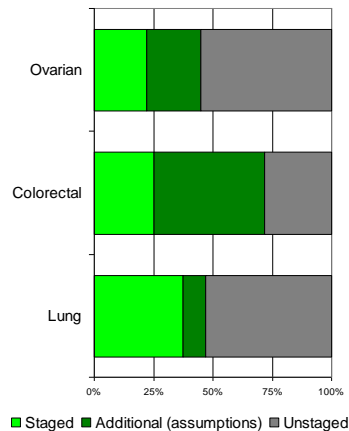
Similar arguments can be made around small tumours, tumours where the nodes are negative etc.

Would be very wary of any statistical prediction of negative mets for a large tumour with positive nodes – fortunately have not seen any!

So can we make assumptions?



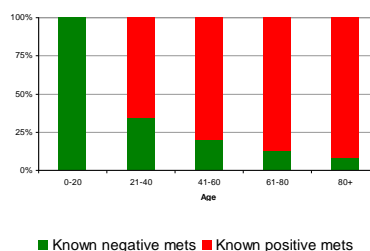
- Method is evidence based
- But is still an art, not a science
 - Look at the four tests and say 'given this, are we comfortable'
- Doesn't make a large difference – improves staging, but doesn't suddenly stage everything
- Saves NHS time to not chase up tumours when it's highly unlikely mets were positive?



What information wasn't used?



- Used data on the tumour and treatment pathway
 - T stage, N stage
 - Surgery, radiotherapy
- Didn't use data on the patient
 - Age, deprivation quintile, ethnicity
 - These data can be good predictors
 - But very cautious to avoid circular results!
- Should we take these factors into account when making assumptions?

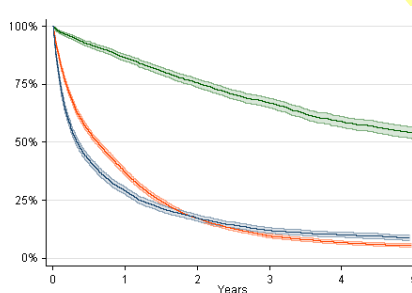


Ovarian cancer by age

What about predicted positive mets?



- Feels 'reasonable' that if nothing is found, people might not write it down
- But methodology also predicts some unknowns as highly likely to be positive!
 - Unknowns look *just* like the ones with positive mets, and 97% of knowns are positive.
- Uncomfortable making such a bold assumption
 - But *can* pass these cases back to Data Quality for rigorous follow-up with Trusts



Late stage colorectal cancers with no surgery and no radiotherapy

Conclusions



- Lack of evidence can provide evidence of lack in limited circumstances:
 - It can be valid to assume that metastases are negative for certain groups of cancers
 - It is clinically inappropriate to do full body scans on all patients
- This methodology can increase the percentage of cancers staged...
...but will not magically stage all cancers
- There are no assumptions so rigorous that they won't be wrong sometimes...
...although the same could be said about collected and inputted data!
- Sophisticated statistical methodologies for dealing with unknowns should not distract from the best way to improve staging data:

Improve data flows from the Trusts in the first place!