Quantitative MRI: the Past, Principles and the Future

Paul Tofts

Emeritus Professor University of Sussex, Brighton, UK
Formerly Chair in Imaging Physics, Brighton and Sussex Medical School , and
Professor at UCL Institute of Neurology, Queen Square, London, UK

www.paul-tofts-phd.org.uk www.qmri.org

Four questions

1. Why should we **Quantify**?
2. Why are **repeatability** and **reproducibility** important?
3. What is a **Perfect qMRI Machine?**
4. What is the proposed **MRI medal system**?

Quantitative MRI: three parts

- 1. the Past

  -- discussion –

- 2. Principles

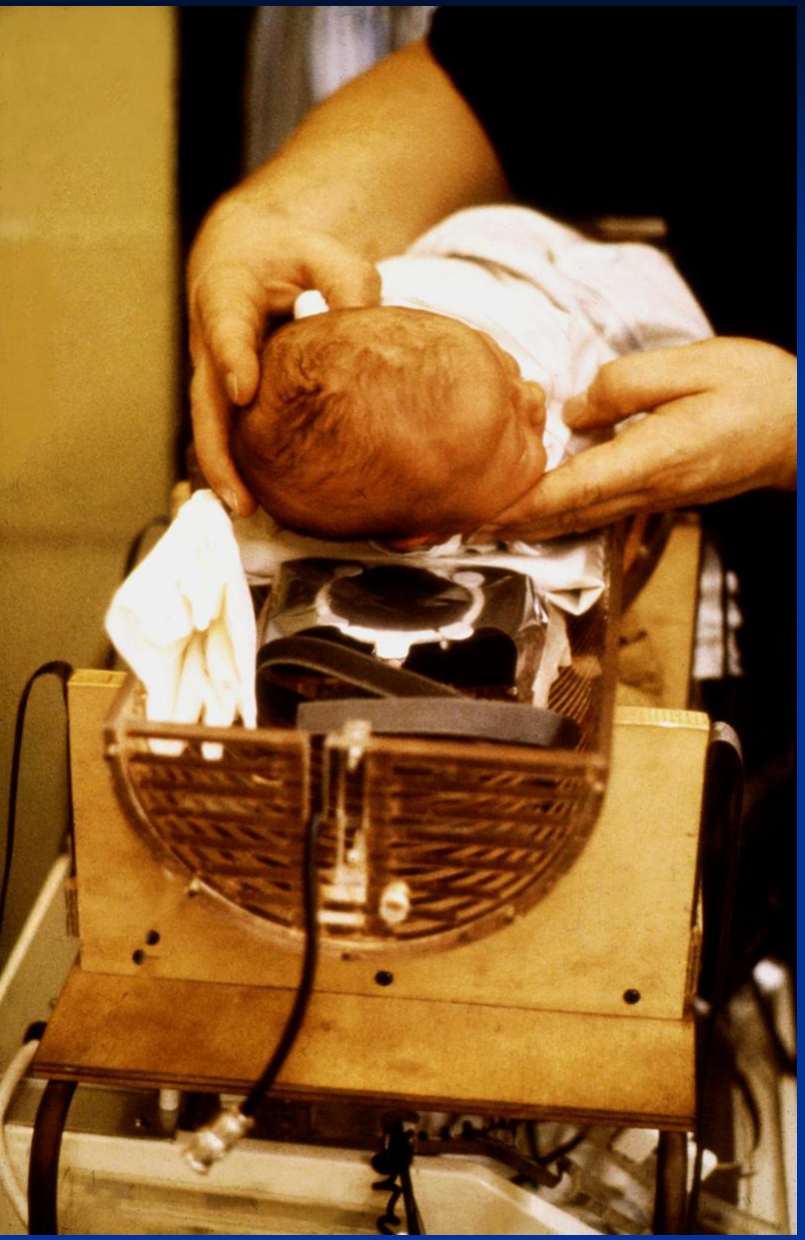
  -- discussion –

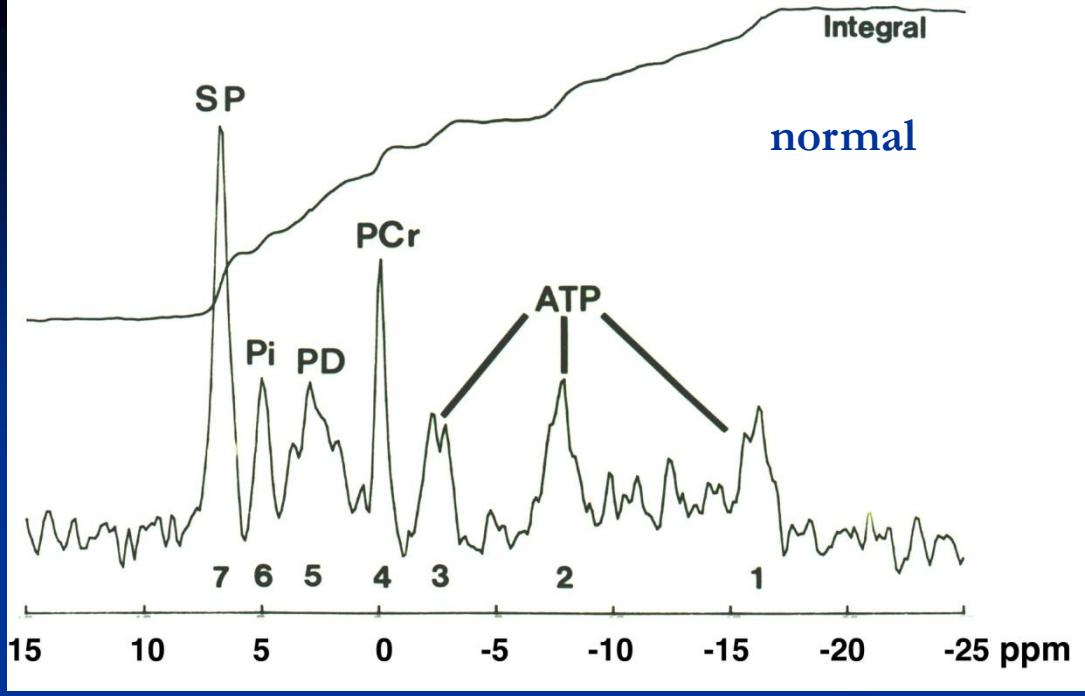
- 3. Future

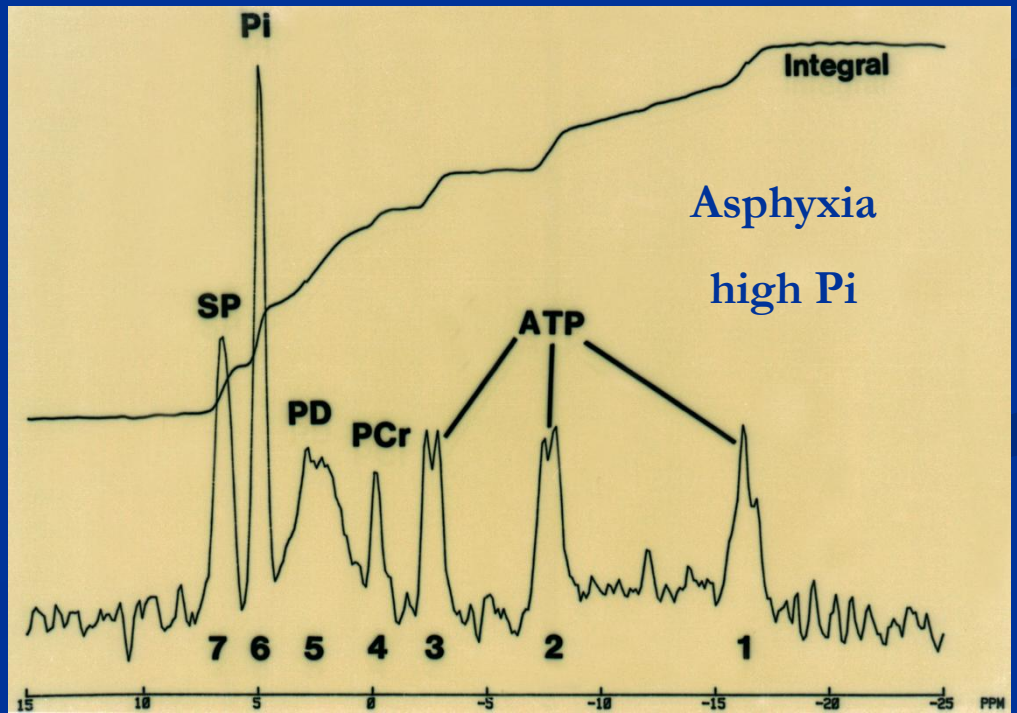
  -- discussion--

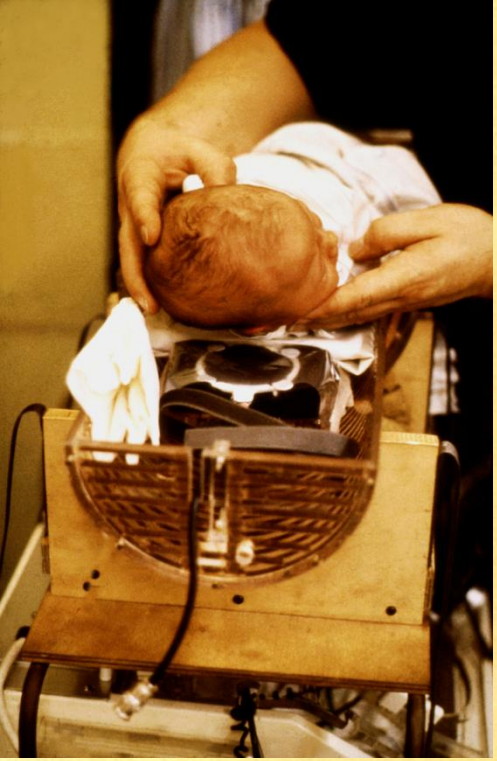
1. Quantitative MRI: the Past

a. ^{31}P MRS in neonates
b. DCE-MRI Gd gives endothelium transfer constant
c. qMT bound protons show myelin
d. MTR histogram predicts clinical score
e. multi-centre – MAGNIMS
f. consensus papers
g. unnoticed qMRI – glioma transformation




Integral
SP
normal
PCr
ATP
Pi
PD
7 6 5 4 3 2 1
15 10 5 0 -5 -10 -15 -20 -25 ppm


Pi
Integral
Asphyxia
high Pi
SP
ATP
PD
PCr
7 6 5 4 3 2 1
PPM



48 hrs
32 hrs
24hrs
9 hrs
15
10
5
0
-5
-10
-15
-20
-25
PPM

^{31}P concentration

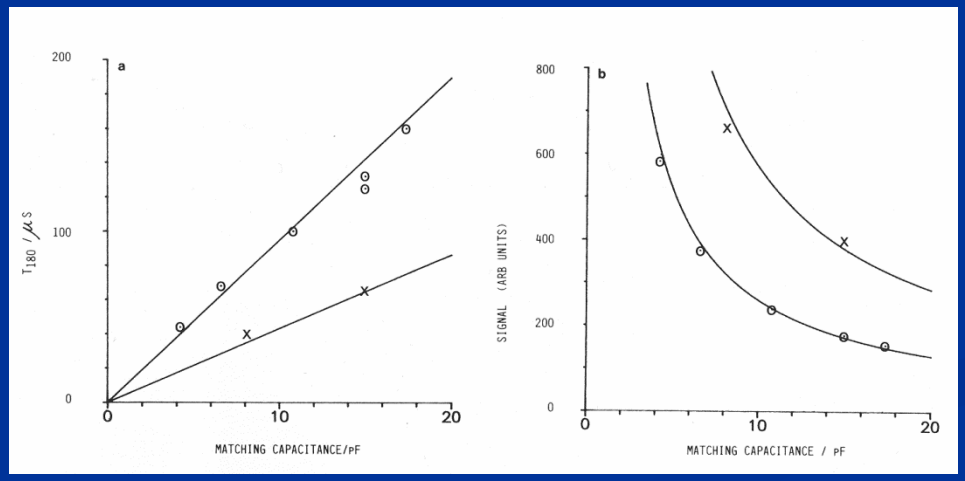
B28

REVIEW PAPER

Dr PAUL TOFTS
Reader in Medical Physics
Institute of Neurology
Queen Square
London WC1N 3BG
Tel: 071-837-3611 ext 4301

A Critical Assessment of Methods of Measuring Metabolite Concentrations by NMR Spectroscopy†

Paul S. Tofts and Susan Wray‡
Institute of Neurology, Queen Square, London, WC1N 3BG, UK, and Physiology Department, University College London, Gower Street, London WC1E 6BT, UK



28

DR. PAUL TOFTS
LECTURER IN MEDICAL PHYSICS
INSTITUTE OF NEUROLOGY
QUEEN SQUARE
LONDON WC1N 3BG
TEL: 01-837-3611

Vol. 1 No. 1 February 1988

NiB

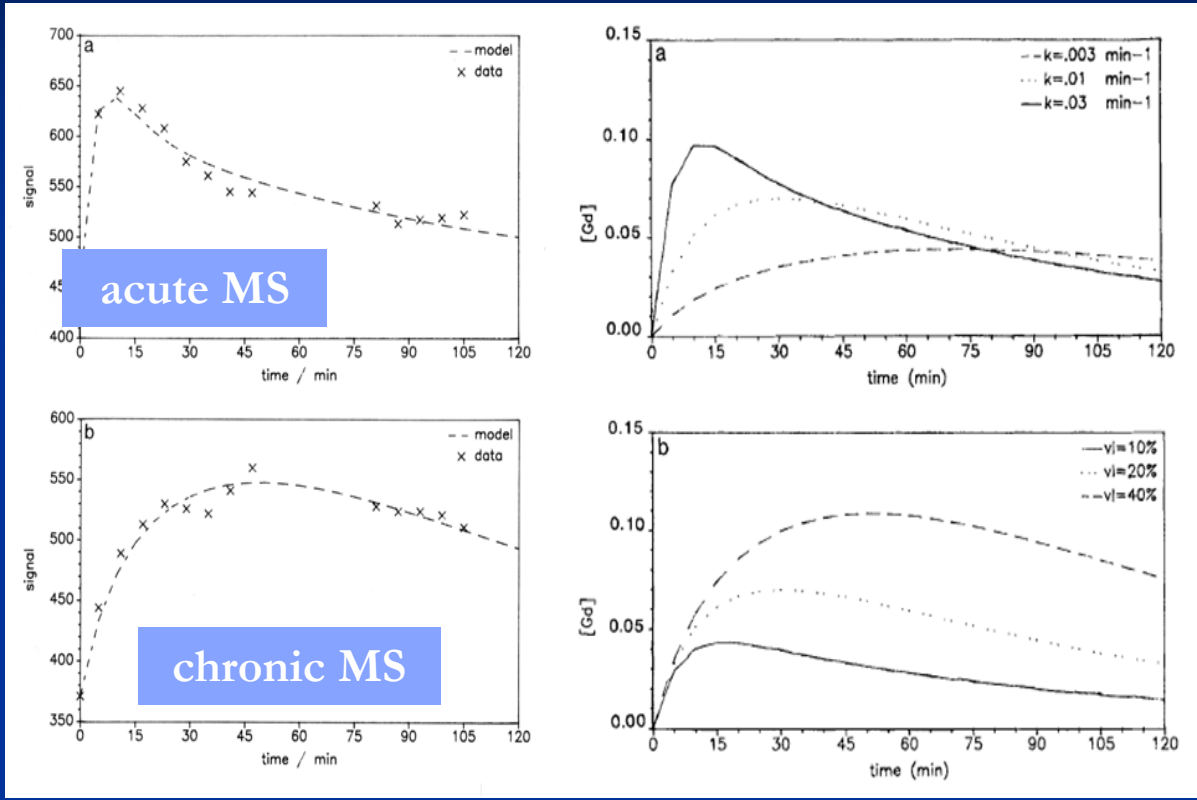
FIRST ISSUE

NMR *in* BIOMEDICINE

AN INTERNATIONAL JOURNAL

DEVOTED TO THE DEVELOPMENT AND APPLICATION OF NMR SPECTROSCOPY *IN VIVO*

Dynamic Contrast-Enhanced MRI



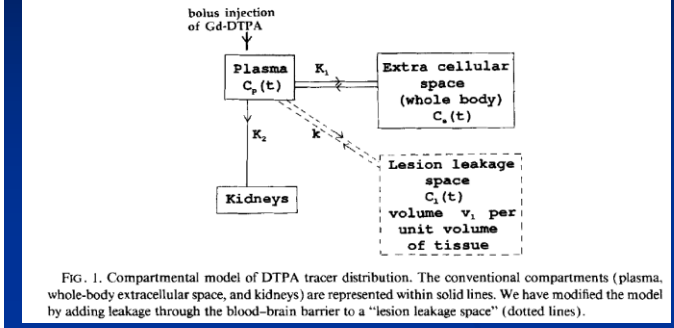


FIG. 1. Compartmental model of DTPA tracer distribution. The conventional compartments (plasma, whole-body extracellular space, and kidneys) are represented within solid lines. We have modified the model by adding leakage through the blood–brain barrier to a "lesion leakage space" (dotted lines).

$$v_l \frac{dC_l}{dt} = k(C_p - C_l)$$

v_l is the size of the Extravascular Extracellular Space

k the transfer constant (depends on permeability and blood flow)

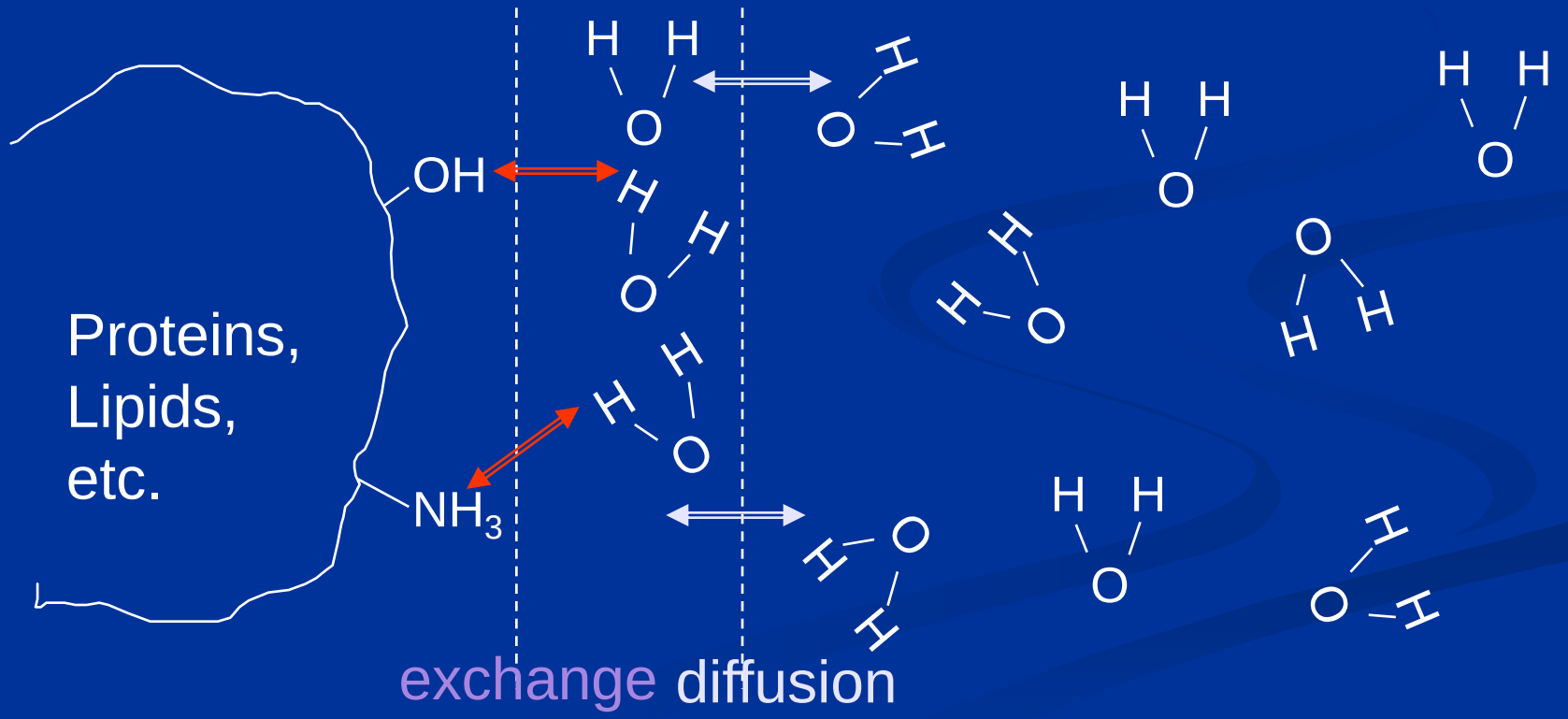
most applications in cancer

Magnetisation Transfer

Macro molecules (invisible)
(bound protons - short T_2)

Surface

Bulk water (visible)
(free protons- long T_2)



qMT in MS

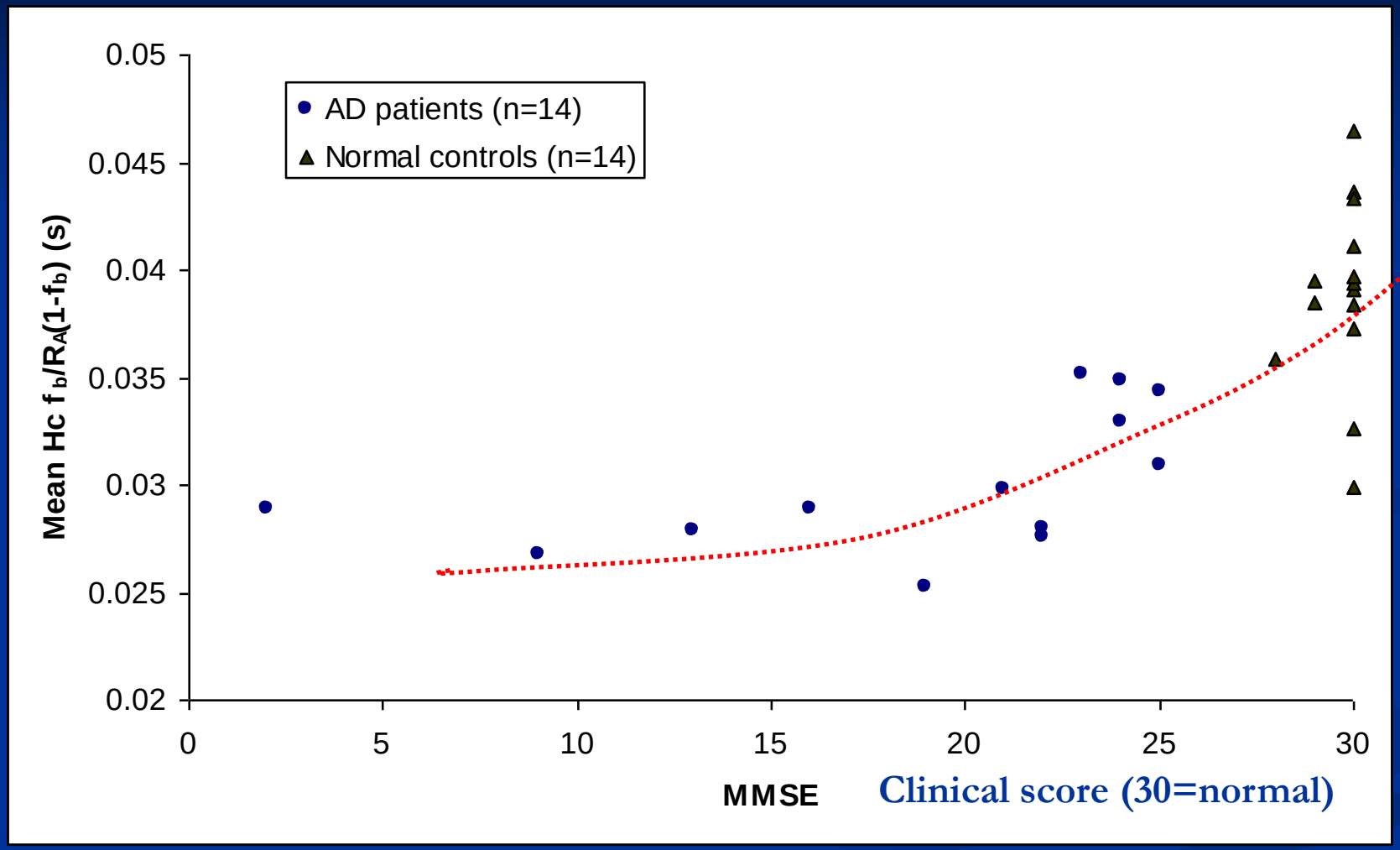
Frontal WM	f_b (%)	p
Control	9.8	
NAWM	8.6	<0.01
Lesion	4.6	<0.01

f_b = fraction of protons that are bound

≈ myelin concentration

Davies et al Mult Scler 2004; 10:607

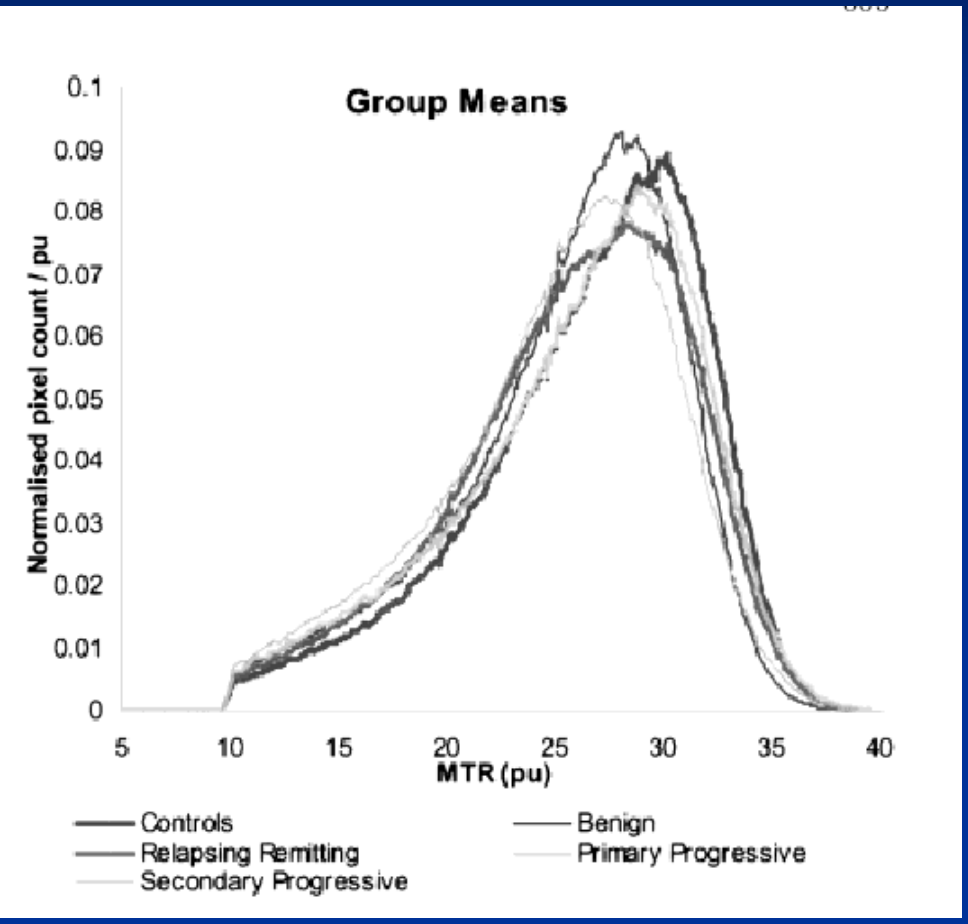
Alzheimer's disease

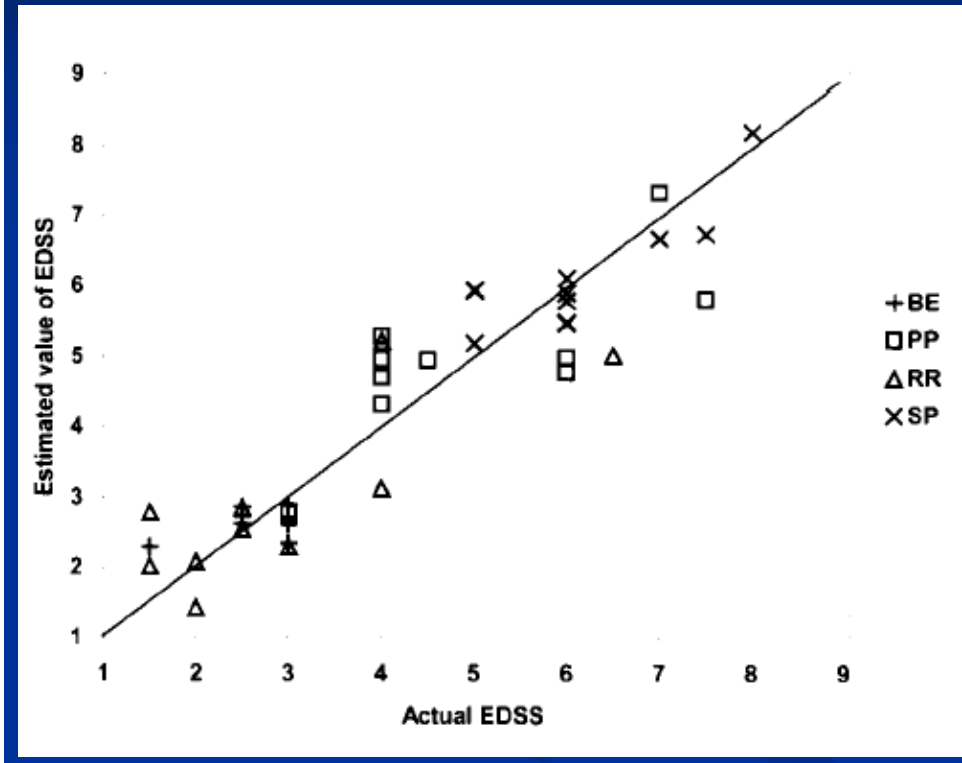


Hippocampal qMT parameter (~ myelin concentration) vs clinical score

Ridha, Fox, Tofts. Quantitative magnetization transfer imaging in Alzheimer disease Radiology 2007; 244:832

MTR histograms in Multiple Sclerosis





Current clinical score (EDSS) can be predicted from histogram (using principle components analysis - PCA)

Whole-brain histogram depends on MS subtype; sensitive to demyelination

Dehmeshki Tofts Magn Reson Med 2001

Reproducibility across centres

1. Much work on multi-centre studies (e.g. MAGNIMS 1990's)
   a. EU funded MAGNetic resonance Imaging in Multiple Sclerosis)
   b. e.g. T2w-lesion load: 5 EU experts in one room

2. 'Protocol Matching' across different manufacturers using standard clinical sequences
   a. works for simple parameters (T_1, D, MT)
   b. relatively easy to implement on a wide scale

3. Travelling controls, phantoms + post mortem brain

4. Complex parameters (e.g. DCE K^{trans}) are often in a 'black box' and may need 'open source' software run on each maker's machine
   May need Research Agreement for each machine

Between-centre difference can be eliminated

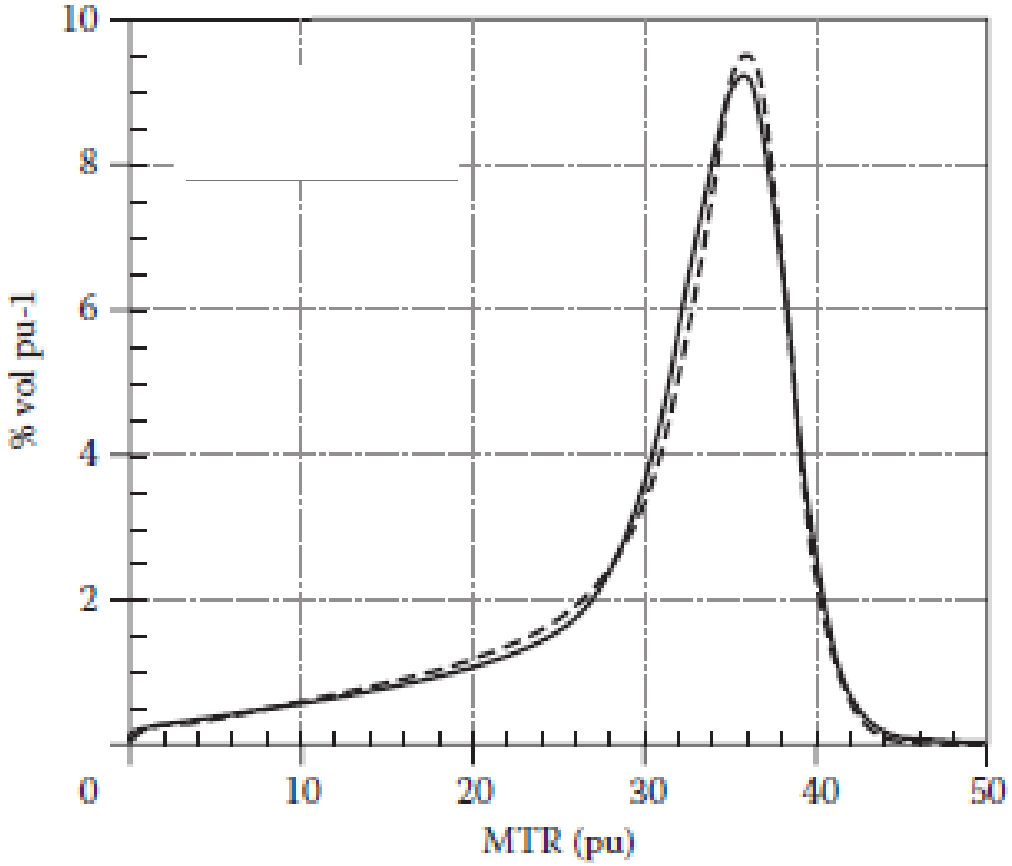


FIGURE 2.9 Matching MTR group histograms from two centres with 1.5T scanners from different manufacturers. By using body coil excitation and standardised histogram generation, inter-centre differences were eliminated. (From Tofts, P.S., *et al.*, *Magma*, 19(4), 209–222, 2006.)

Consensus papers

JOURNAL OF MAGNETIC RESONANCE IMAGING 10:223–232 (1999)

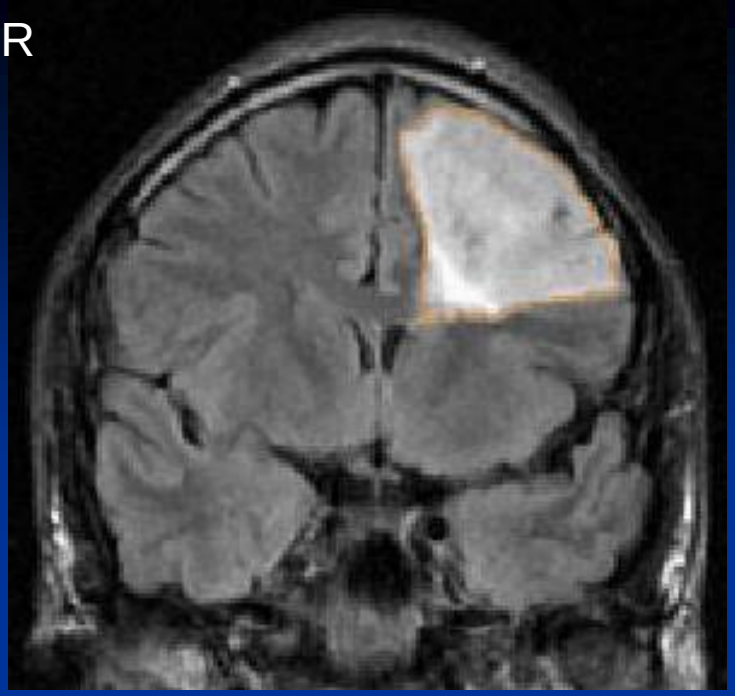
Review

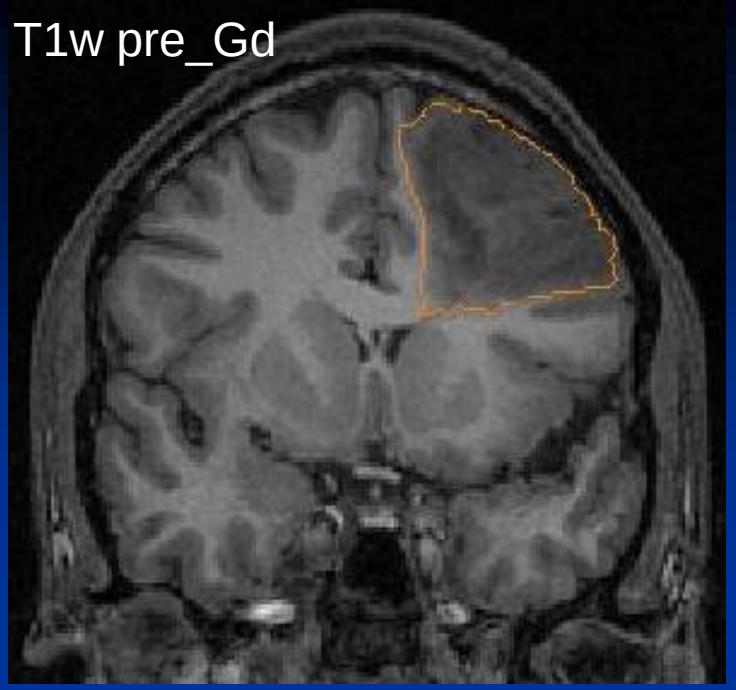
Estimating Kinetic Parameters From Dynamic Contrast-Enhanced T_1-Weighted MRI of a Diffusable Tracer: Standardized Quantities and Symbols

Paul S. Tofts, DPhil,[1*] Gunnar Brix, PhD,[2] David L. Buckley, PhD,[3] Jeffrey L. Evelhoch, PhD,[4] Elizabeth Henderson,[5] Michael V. Knopp, MD,[6] Henrik B.W. Larsson, MD,[7] Ting-Yim Lee, PhD,[5] Nina A. Mayr, MD,[8] Geoffrey J.M. Parker, PhD,[1] Ruediger E. Port, MD,[6] June Taylor, PhD,[9] and Robert M. Weisskoff, PhD[10]

1. Identify leaders
2. Invite them to a meeting
3. Write the paper (on methodology, analysis, terminology)
4. Reject any papers that do not use this consensus

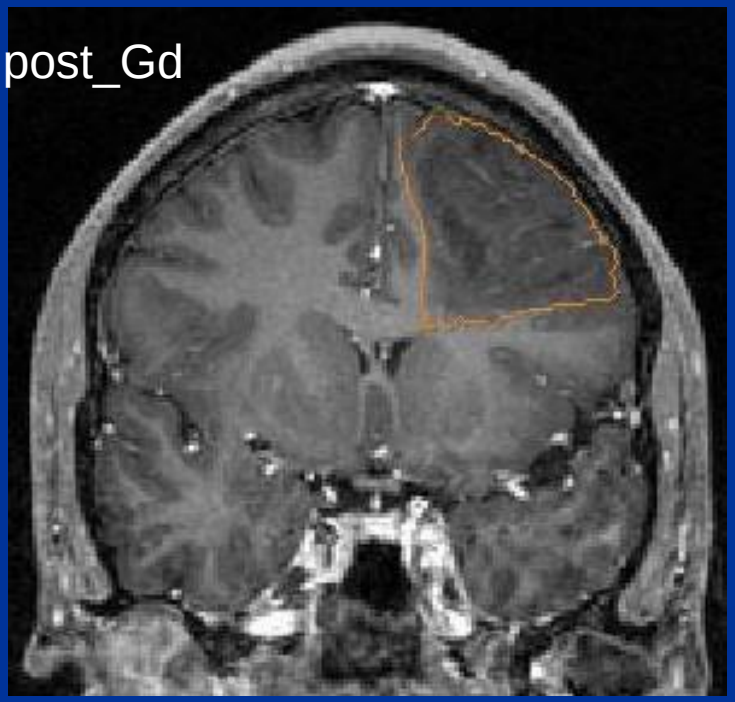
2003 citations (October 2019)



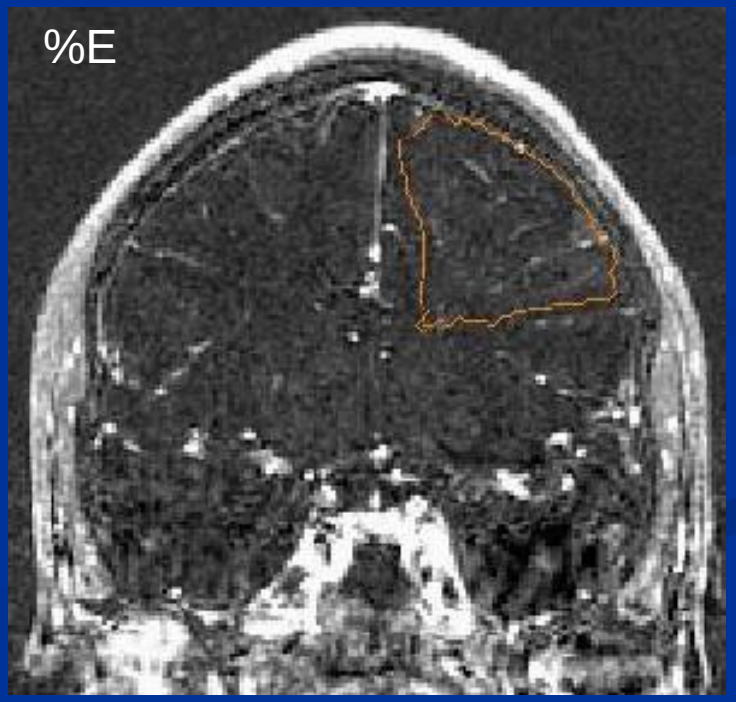


Low Grade Glioma

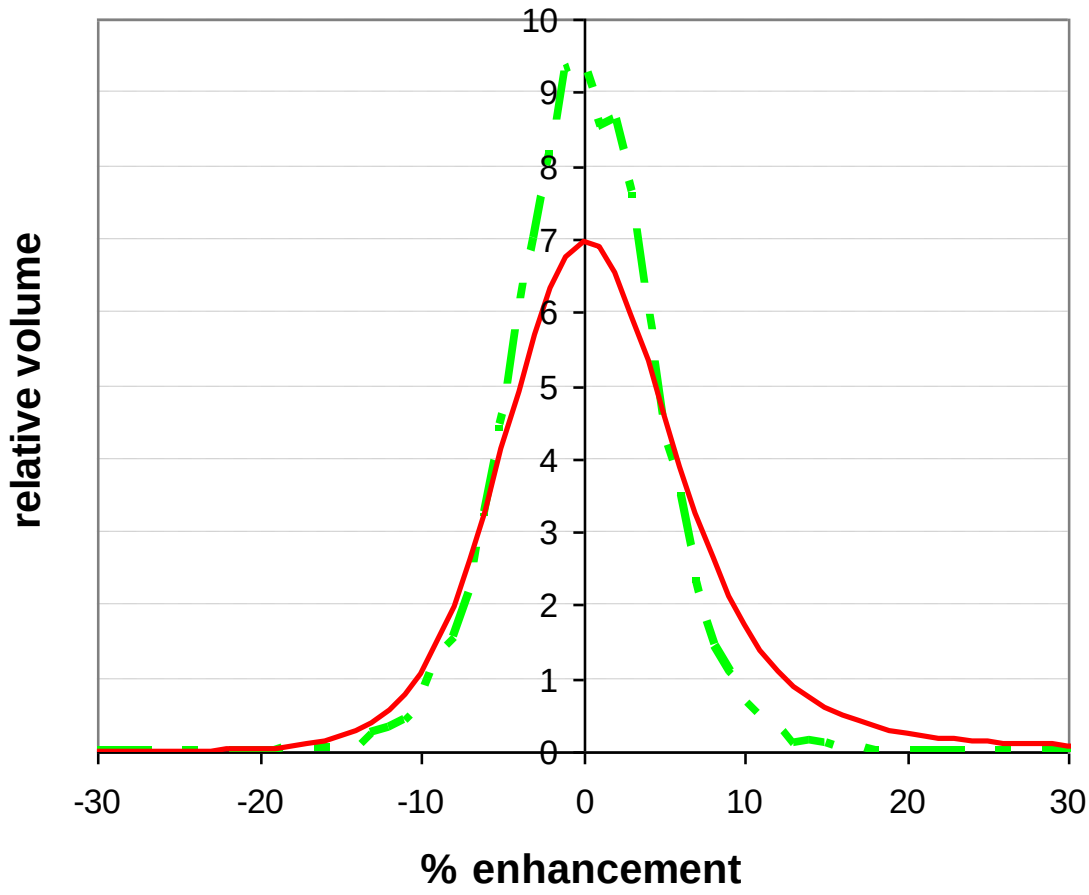
'not visibly enhancing'



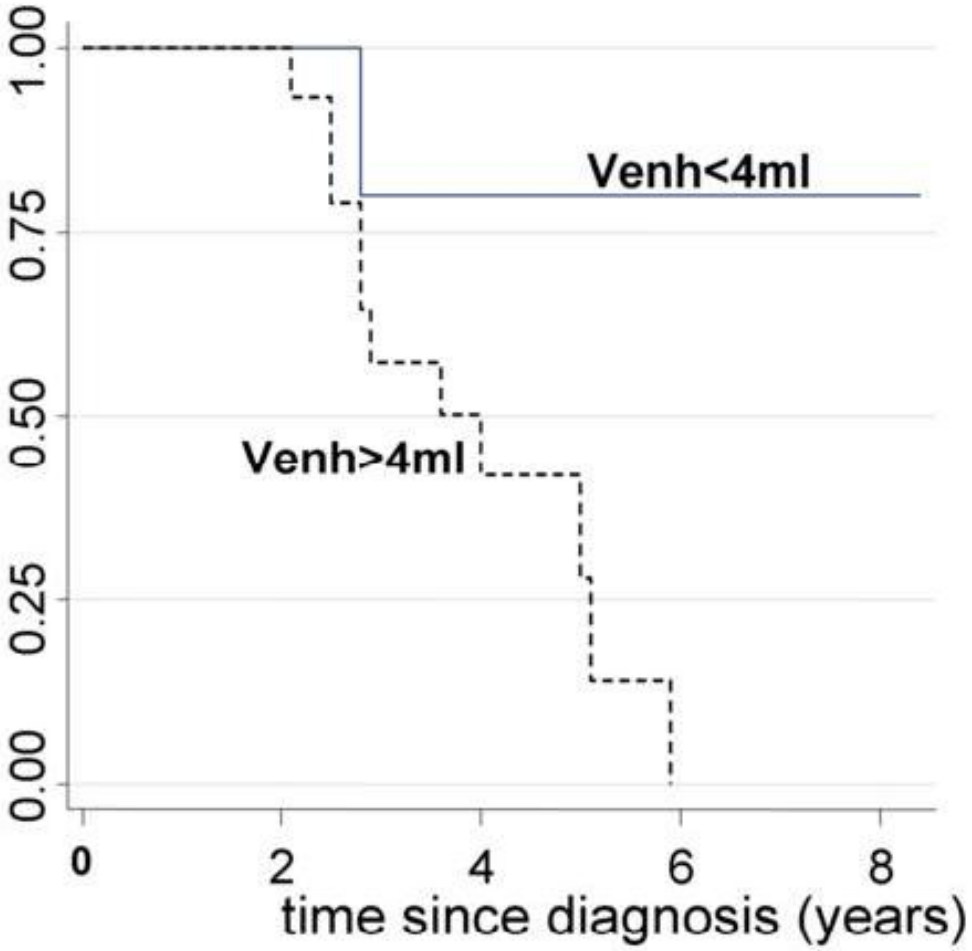
c



d



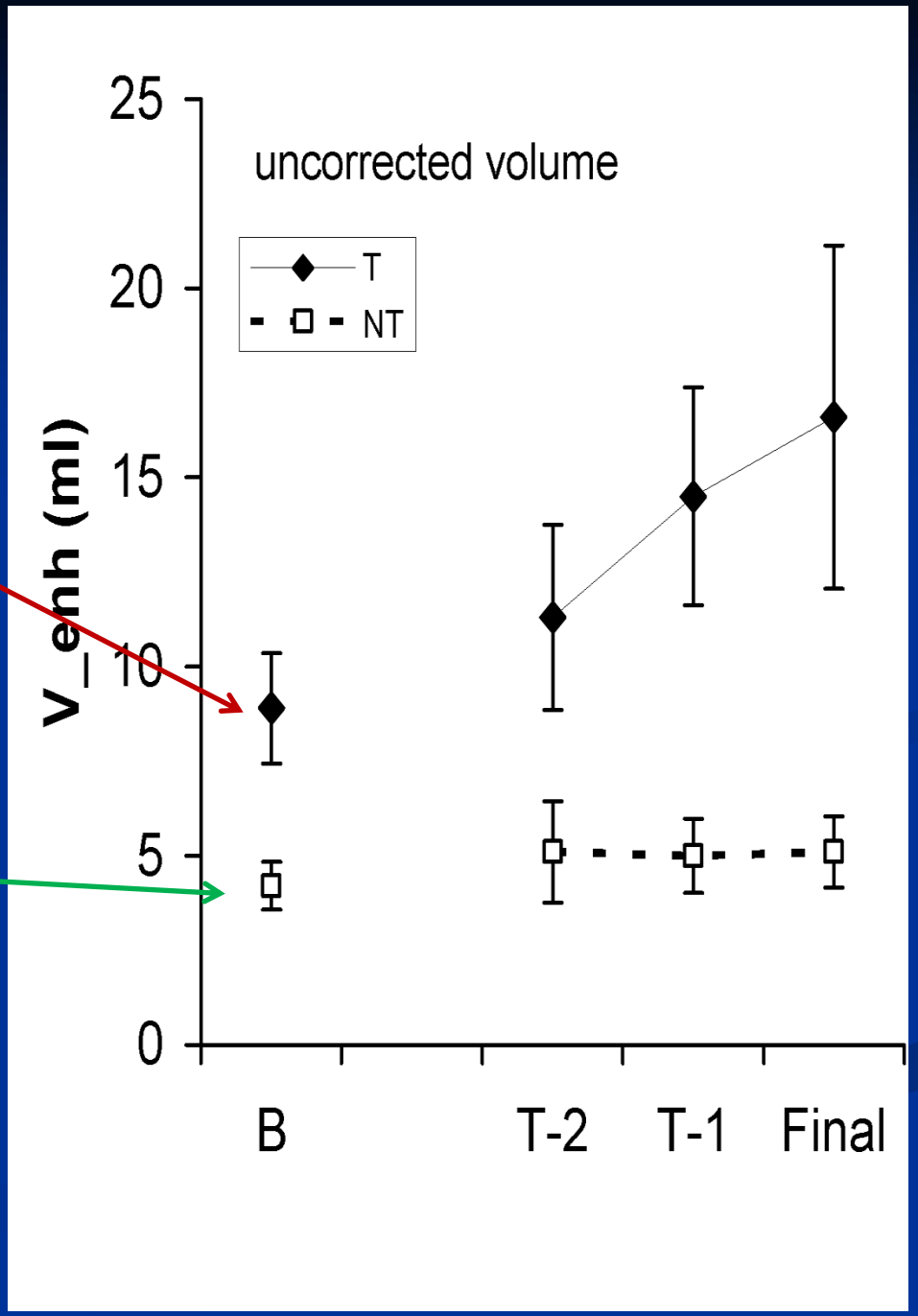
Measure size of RHS tail = volume of abnormal tissue



Kaplan-Meier survival plot, using uncorrected volume from baseline scan

p<0.039 at 5 years

- Transformers show progressive increase in enhancing volume
  - different from NT even at baseline
- Non transformers are stable
  - small SD; homogeneous group



Tofts JMRI 2007; 25:208-14
Quantitative analysis of whole-tumor Gd enhancement histograms predicts malignant transformation in low-grade gliomas.

1. Quantitative MRI: the Past

a. ^{31}P MRS in neonates
b. DCE-MRI Gd gives endothelium transfer constant
c. qMT bound protons show myelin
d. MTR histogram predicts clinical score
e. multi-centre – MAGNIMS
f. consensus papers
g. unnoticed qMRI – glioma transformation

2. Quantitative MRI: Principles

a. why quantify?
b. the Books
c. accuracy & why Random Error is the Enemy
d. phantoms vs healthy controls
e. data acquisition
f. data analysis
g. Upgrades are also an Enemy
h. Statistics are friends

What is qMRI?

Quantification = measure

Quantity

e.g. body mass

reliable, accurate, reproducible, easy

Quantification

- Quantify – to measure a quantity (size, weight, blood sugar, cholesterol …)
- Medical images have been *qualitative*
  - **Look**; human assessment; experience needed
- Imaging is becoming *quantitative*
  - **Measure** e.g. tumour size, water content, tissue destruction, volume of MS lesions…

Why is qMR needed?

1. Measurement concepts - sources of variation
2. Specificity - new biological quantities
3. Scientific instrument following long tradition of measurement in astronomy, physics, chemistry, electrical engineering…
4. Measure subtle 'invisible' changes ; diffuse or small, in 'Normal-Appearing' brain tissue

Psychometric measures
desirable properties

- Sensitivity
  - does the quantity alter with disease?
- Validity
  - Is it relevant to the biology?
- Reliability
  - Is it reproducible?

qMRI of the brain, 1st edition p68

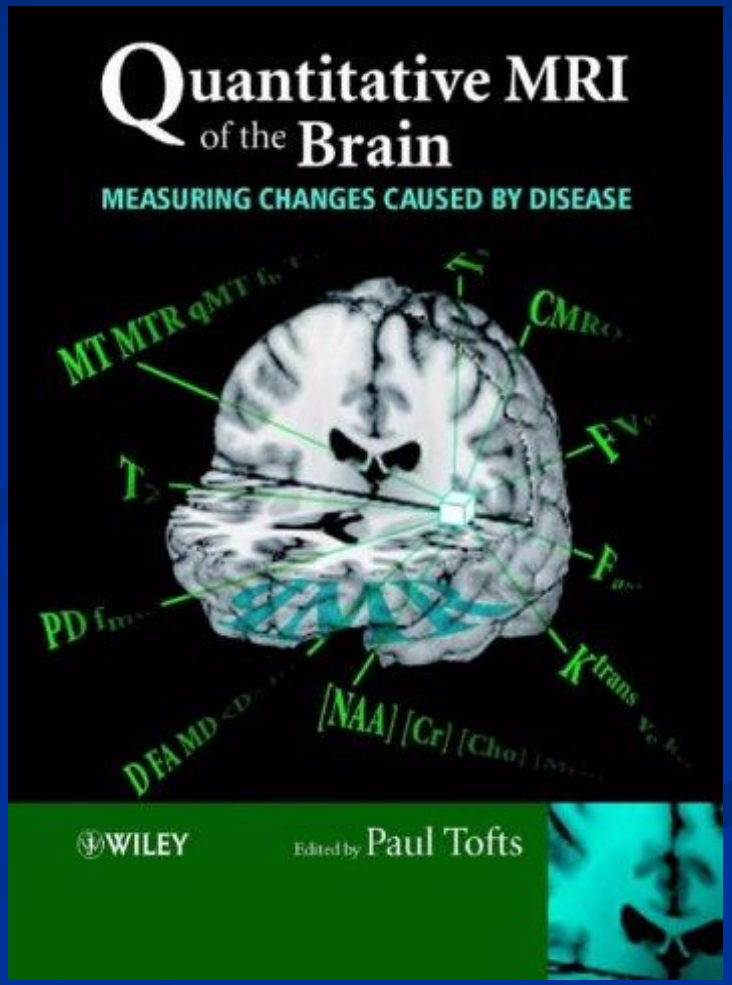
qMRI – a technology whose time has come

Medical Imaging

meets

Measurement Science

British Medical Association Radiology book prize 2004



'The pre-eminent role of imaging now requires a new level of metric - quantitative measurements'

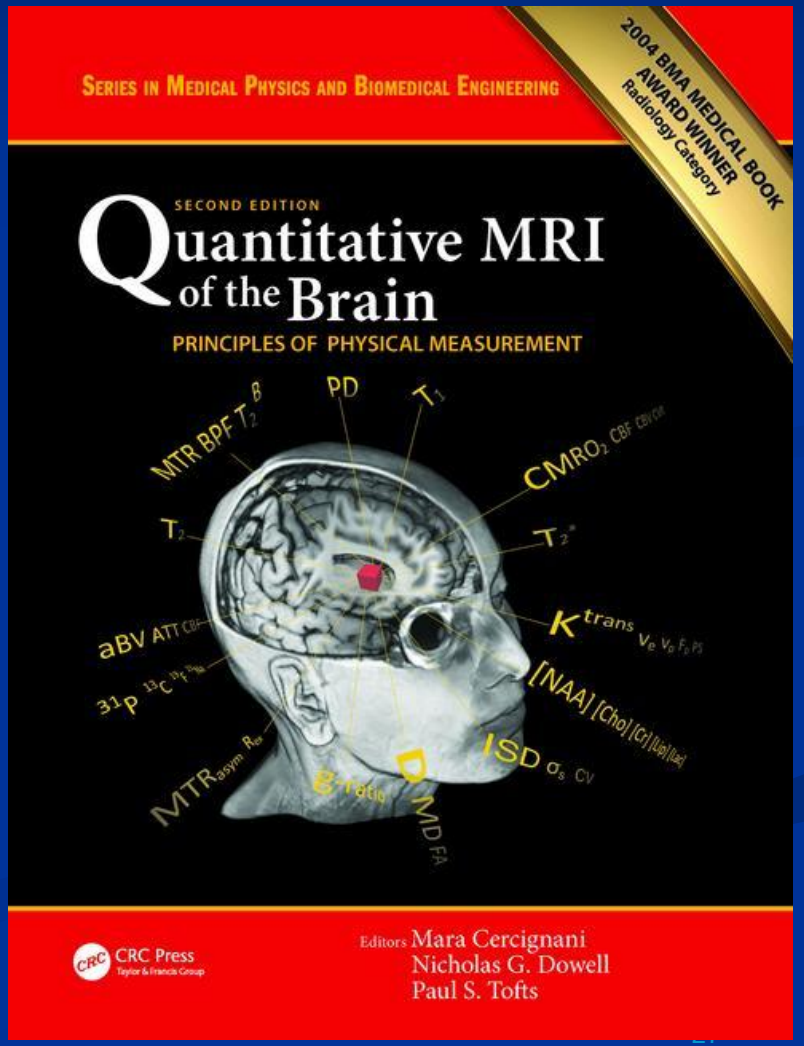
Robert I Grossman MD, Chair of Radiology, New York University

new edition 2018

€120 hardback; €50 eBook (Amazon or CRC press)

see qmri.org (some author pre-prints)

Contents



UK Institute of Physics and Engineering in Medicine
Report 112 2017

Is accuracy important?

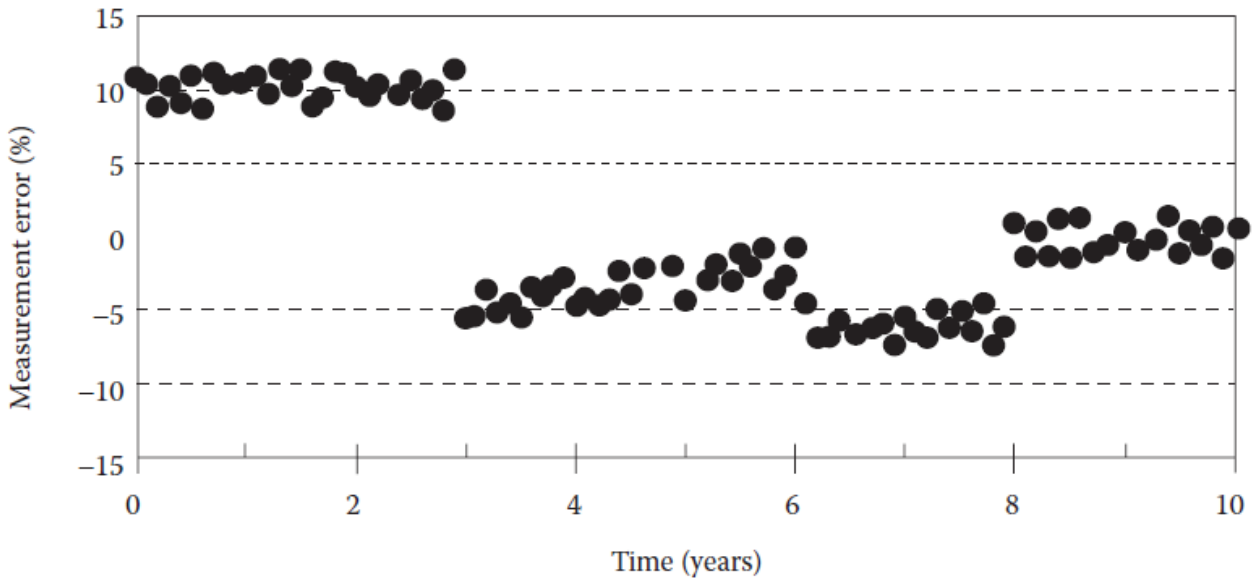


FIGURE 3.3 Long-term precision is dominated by instability in the systematic error. Simulation of fictional change in measurement error over time, during a longitudinal study. Short-term precision is good, and a study completed in the first 3 years is unaffected by the large systematic error (i.e. poor accuracy). A major upgrade at Year 3 dramatically changes the systematic error. A subtle drift in values takes place, followed by two more step changes, at the times of operator change and a minor upgrade. At Year 8 the sources of systematic error are finally identified and removed, giving a system that should provide good accuracy and hence long-term precision for many years.

Is accuracy important?

In a single centre short study – probably not

In longer studies – yes (withstand upgrades)

In multi-centre studies – yes

(unless you can replicate the sources of inaccuracy at each site – 'protocol matching')

Why does random error matter?

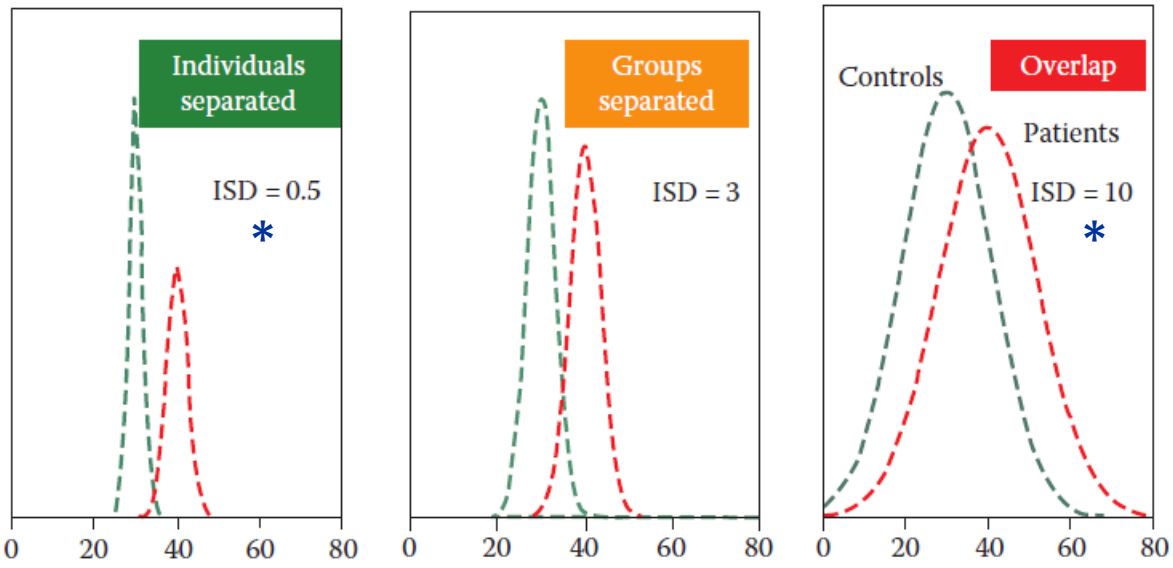


FIGURE 3.5A Simulation showing how magnitude of ISD affects ability to use an MR parameter to separate groups and individuals. Group separation is 10 units. With ISD = 10 (right-hand image), the groups overlap, and considerable statistical power would be needed to separate them (see Chapter 1, Figure 1.3). A reduced ISD = 3 (centre) gives a good group separation c) a further reduction to ISD = 0.5 (left-hand image) enables individuals to be accurately classified into their group.

* ISD = Instrumental Standard Deviation (repeatability)

Why is repeatability important?

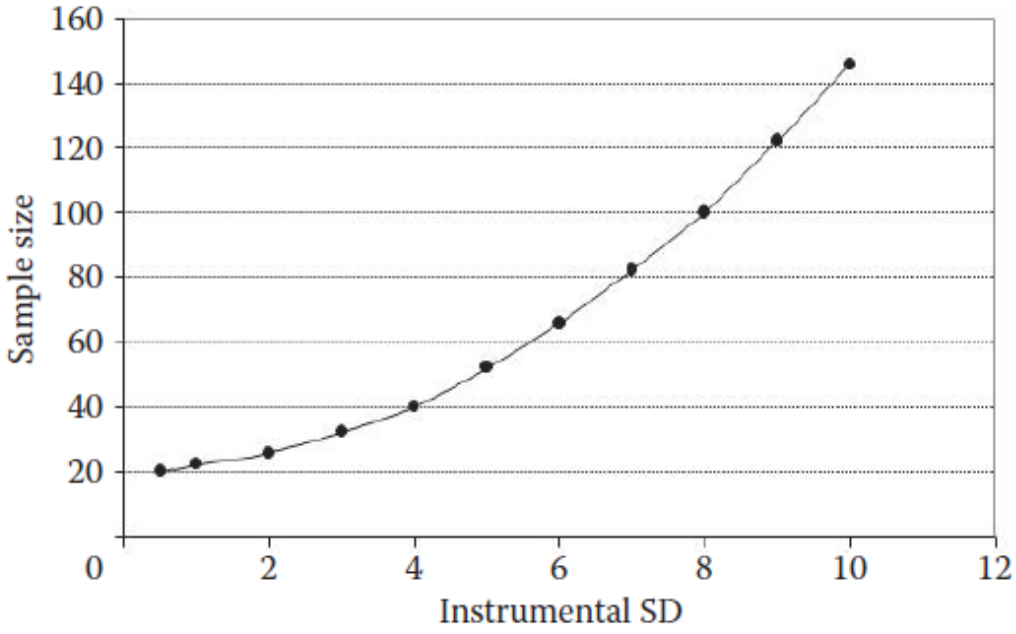


FIGURE 1.3 The effect of instrumental precision (ISD) on the power of a study and the required sample size. By reducing the ISD, the sample size required is dramatically reduced, with a consequent saving in the cost and duration of the study. This is a simulation based on group comparison between controls (parameter value mean = 100, SD = 3) and the same number of patients (effect size = 5, SD = 4.25). Power P = 80%, significance level $\alpha = 0.05$, using G*power3 which is an established software that can be downloaded free of charge.

Phantoms and healthy Controls for QA

TABLE 3.1 Relative Advantages of Phantoms and Healthy Controls for Quantitative Quality Assurance

	Simple Phantom (Test Object)	Healthy Control Subjects
Availability[b]	Good	Reasonable
Accuracy	Potentially good (e.g. volume)	True value unknown[a]
Uniformity	Poor in gels, good in liquids	Good in white matter
Temperature dependence	D,T_1,T_2 change 2%–3%/^{O}C	Homeostatic temperature control
Stability	Potentially good (e.g. volume) but can be unstable (e.g. gels)	Usually stable
Realism	Generally poor; *in vivo* changes cannot be realistically modelled; B_1 distribution different	Good but no pathology
Standard design for multicentre studies?	Can be made	Use normal range, or travelling subject(s)

[a] Although normal values have a narrow range – see Table 3.5.
[b] Though see institutional constraints (Section 3.5.1).

Normal ranges: T_1 4-6%; MD: 3-5%; MTR: 1-2%

With correction for age etc, and control of ISD these would probably be reduced

Phantoms vs controls

1. Good phantom performance necessary but not sufficient
2. Phantom – beware RF dielectric resonance
3. Multi-centre: travelling phantoms? Controls?

Why are physicist so interested in scanning normals?

- Repeatedly!
  - Understand and minimise all the sources of variation
- Serial study
- Cross-sectional study
- Influence of instrumental variation on sample size in power calculation

The Measurement Process: MR Data Collection and Image Analysis

Contents

Some quantities depend on acquisition parameters (e.g. T_2, MD depend on TE's)

What causes random variation?

TABLE 3.3 Potential Sources of Error in the MRI Measurement Process[a]

	Random Error	Systematic Error
Biology	Normal variation in physiology	
Data collection	Position of subject in head coil	B_1 error
	Coil loading (corrected by prescan?)	Slice profile
	Prescan procedure setting B_1	K-space sampling (in FSE, EPI)
	Position of slices in head	Partial volume
		Operator training
	Gd injection procedure	Software upgrade
	Patient movement (cardiac pulsation)	Hardware upgrade
	Patient movement (macroscopic)	
	Image noise	
	Temperature (phantoms only)	
Image analysis	ROI creation and placement	Operator training
		Software upgrade

Note: In their simplest forms, random error is associated with short-term unpredictable variation, whilst systematic error is fixed. However some random processes (e.g. positioning) might only show up over a longer time scale (caused e.g. by change of radiographer [technician]), whilst some sources of systematic error might vary with time (e.g. operator training). ROI = region of interest.

B_1 errors

Quality of parameter estimates depends on quality of acquisition

B_1 and image noise often dominate parameter uncertainty

[poor acquisition cannot be fixed by post-processing!]

1% error in flip angle FA gives 2% error in T_1 (in Variable Flip Angle method)

Slice selection is bad news – use 3D acquisition?

Optimisation of acquisition procedure

Minimising the effect of image noise

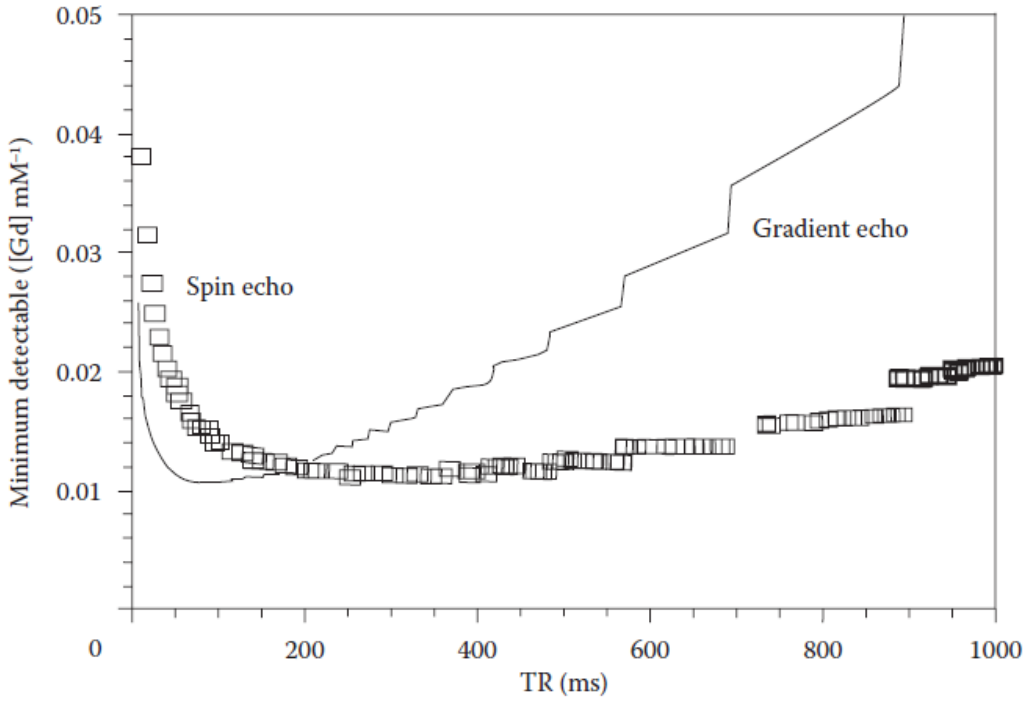


FIGURE 2.5 Sequence optimisation by noise modelling. Mathematical modelling of image noise propagation predicts the minimum amount of Gd contrast agent that can be detected using a T_1-weighted sequence. By optimizing the repetition time, *TR*, in a spin echo or gradient echo sequence, its performance in white matter can be optimised. Theory indicates that, for a spin echo, the optimum $TR = T_1/2$ (here it was assumed $T_1 = 600$ ms). The gradient echo (flip angle, FA = 50°) can achieve the same sensitivity, provided the correct *TR* is used. The examination time was fixed at 10 min. (From Tofts, P.S., *Magn. Reson. Imaging*, 14(4), 373–380, 1996.)

Image data analysis

- Region of interest
  - Test a specific location (prior information and hypothesis)
- Histogram
  - Whole brain; unbiased; for diffuse disease
- Voxel-Based Morphometry VBM
  - Unbiased testing of many locations
  - Each location can be correlated with external score (clinical, genetic, proteomic, cognitive)
- Texture
  - 'dirty white matter'
  - tissue often becomes more heterogeneous in disease

Upgrades are also an Enemy

- Any long-term study needs stability
- Any serious change will need repeated validation of qMR method
- Changes can be software, hardware, field strength
- Many quantities ought to remain unchanged with good methodology (e.g. volume)

Statistics are friends

- In a group comparison study, often group differences are reported as p-values
- If no significant difference seen, was this because:
  - There is no biological difference between the groups
  - The instrumentation is rubbish (large instrumental SD: ISD)
- Better: give confidence limits for group difference, measured group SD, and estimated ISD
- Then studies can be evaluated, compared and pooled

2. Quantitative MRI: Principles

a. why quantify?
b. the Books
c. accuracy & why Random Error is the Enemy
d. phantoms vs healthy controls
e. data acquisition
f. data analysis
g. Upgrades are also an Enemy
h. Statistics are friends

3. Quantitative MRI: the Future

a. why are we here??
b. The Perfect Machine
c. Medals
d. Understanding Normality
e. Understanding machine variation
f. Why is qMRI not like a thermometer?
g. Resources at qMRI.org

why are we here??

- No-one ever wished on their death bed that they had spent more time in the office (from a time-management course)
- One of the 10 keys to happiness is to do **meaningful work**
  (from Action for Happiness)
- Break out of continual re-implementation of methods

Perfection is possible

A Perfect Quantitative MRI machine is one that, in making a measurement, contributes no significant extra variation to that which already exists from biological variation.

The concept of the 'Perfect Machine' originates in the building of the 200 inch Palomar telescope in 1933-48.

inspiration: In Thomas Mann's *Death in Venice*, the writer is on the Venice beach. He sees the detail, in the foreground: children constructing a sand castle. He turns his gaze to the horizon, empty and infinite. What would it be to be a measurement hero?

From ***Quantitative MRI of the Brain*** p10

Medals for Perfection

TABLE 1.3 qMRI Medals for Perfect Machines: A Proposal

Medal	Target Study	Criterion	Motivation
Bronze	Group comparison	ISD < 0.3 GSD	(a)
Silver	Multicentre study	BCSD < GSD	(b)
Gold	Serial study	ISD < 0.3 WSSD	(c)

Note: SD: standard deviation; BSD: biological SD; GSD: group SD; ISD: instrumental SD; ICSD: inter-centre SD; BCSD: between-centre SD; WSSD: within-subject SD.

[a] In a group comparison, within-group variation GSD^2 should dominate (i.e. machine variation ISD makes an insignificant contribution to total within-group variation).

[b] The effect of between-centre variation (BCSD) should be less than within-group variation.

[c] In a serial study, total within-subject variation $WSSD^2$ should dominate (i.e. machine variation ISD makes an insignificant contribution to total within-subject variation).

NB A medal could exist for each qMR parameter.
Inspiration: the lifetime work of John Harrison, who constructed stable travelling clocks. The Longitude prize of £20k was offered by the British parliament in 1714, in response to loss of life at sea and an urgent need for better navigation. This medal scheme might be attractive to a philanthropist.
from *Quantitative MRI of the Brain* p10

Normality: normal range depends on repeatability

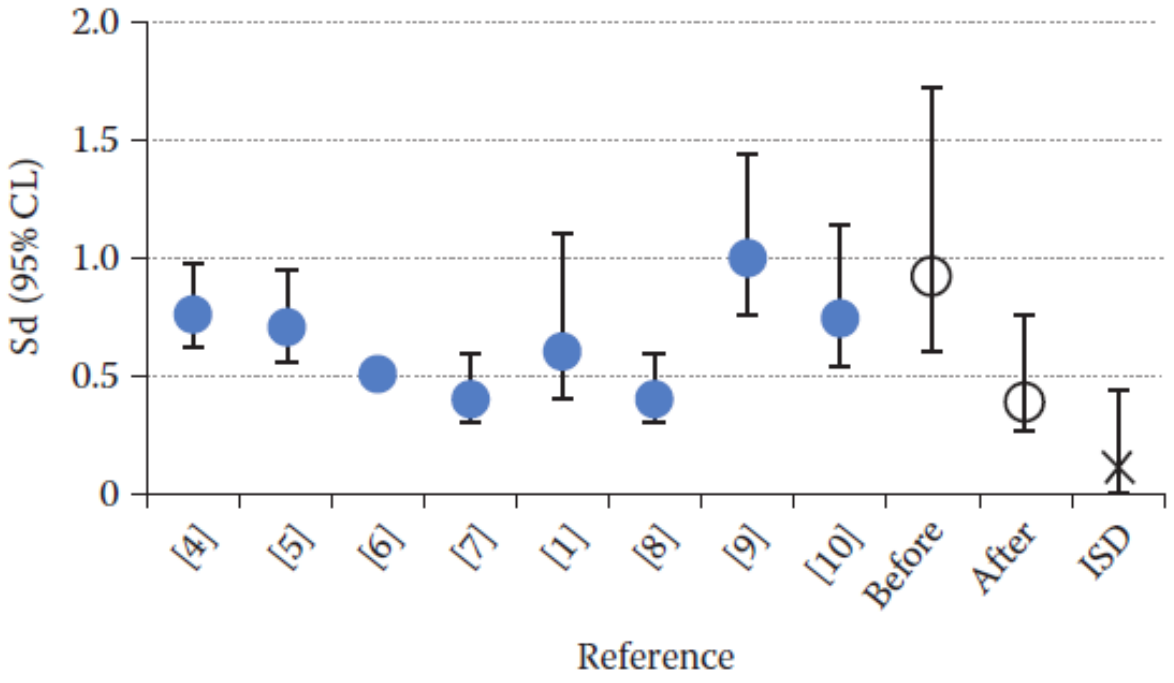


FIGURE 3.8 Normal variation for white matter MTR, and influence of ISD. Blue circles are published values of SD (units for MTR are pu; mean was 38–40 pu) from eight centres; error bars show uncertainty in sd estimate (Equation 3.2). *Before* is authors' first value, almost the highest value of nine centres. After solving a scanner instability problem (Figure [stability] in Chapter 2), ISD was low (≃0.2 pu) and the re-measured normal range (*after)* dropped to the lowest value of nine centres.

An invisible problem

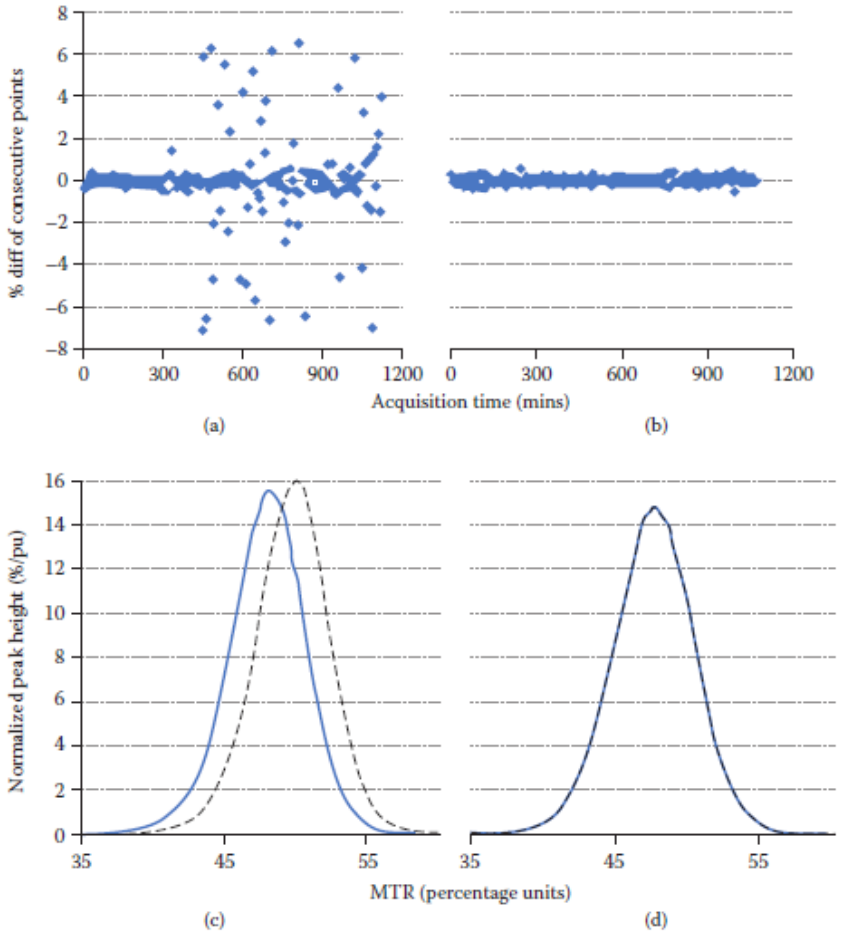


FIGURE 2.8 Unsuspected scanner instability – an invisible problem. MTR histograms showed a large within-subject variation (c). Repeated scanning of a phantom overnight showed large random variation (a). After changing transmitter boards, the scanner was stable (b) and MTR histograms were reproducible (d). (Data from NG Dowell, originally presented in (Haynes *et al.*, 2010) (From Haynes, B.I.,*et al.*, Measuring scan-rescan reliability in quantitative brain imaging reveals instability in an apparently healthy imager and improves statistical power in a clinical study. ISMRM annual scientific meeting, Stockholm, p. 2999, 2010.)

Understanding machine variation

1. More to come

2. Not just image noise

3. Low level 'noise' masking subtle Gd enhancement
   a. Short Term Long-range Fluctuations probably originate from pulsatile movement of the bright Superior Sagittal Sinus (<1%)
   b. Movement through the nonuniform B_1^- receive field, not corrected by registration software.
   c. (ISMRM Paris 2018 poster)

Why is qMRI not like a thermometer!

1. Thermometer (or voltmeter): works, reliable .....
2. qMR from vendors: another story
3. Killer App may drive vendor implementation (MD in stroke, K^{trans})
4. drivers: pharma trials ... NHS treatment decisions

The future

1. Type A and B errors
   a. Papers from NPL and NIST 'estimating uncertainty in measurement'
   b. Random vs systematic error, depends on time scale
   c. for voltage or temperature we just have max uncertainty (95%?)
   d. ADC alkane measurements: propagation of errors in G,T etc
   e. Consensus paper on how to... ?

2. ISMRM reproducibility challenge

3. National Measurement centres: use their expertise and concepts
   NIST – National Institute of Standards and Technology, USA
   NPL – National Physical Laboratory, UK
   PTB – National Metrology Institute of Germany

qMR – the future

qMR is becoming a turn-key application

Happy Snappy MRI Camera

transforming into

Scientific Instrument

We are witnessing

paradigm shift

technological revolution

Link: qmri.org/hack2019 Nikola Stikoff

- ISMRM special workshop; consensus position paper
- Publish specific medals e.g. T_1, MD (some may already exist)