

# La statistica della datazione della Sindone

Gian Marco Rinaldi

© Gian Marco Rinaldi , 2012

- (I) Il resoconto di *Nature*
- (II) Le accuse dei sindonologi
- (III) La qualità delle datazioni negli anni '80

Fra gli svariati tentativi dei sindonologi per screditare la datazione medievale della Sindone, c'è stato chi ha attaccato il risultato in base a un esame dell'analisi statistica che fu condotta in origine e che è pubblicata sul resoconto di *Nature* del 1989 [6]. I sindonologi che se ne sono occupati direttamente sono stati pochi. Quello che ha più insistito è stato il belga Remi Van Haelst che nel corso degli anni è tornato ripetutamente alla carica con diversi articoli. In Italia c'è stato dapprima Ernesto Brunati, poi recentemente la linea di attacco basata sulla statistica è stata rilanciata in un libro di Marco Tosatti, *Inchiesta sulla Sindone* [11].

Dal tenore delle critiche, appare evidente che i sindonologi non hanno capito quali siano stati i metodi di analisi statistica impiegati nel resoconto di *Nature*. Lì si può in parte scusare perché quel resoconto è sintetico e talvolta omette informazioni o spiegazioni che invece sarebbero state necessarie. Nella prima parte di questo articolo provo a illustrare in dettaglio (o talvolta cerco di indovinare) i metodi di analisi di *Nature*. Nella seconda parte passo in rassegna le accuse dei sindonologi. La terza parte è una retrospettiva su quale fosse il livello di qualità delle datazioni negli anni '80.

## I - Il resoconto di *Nature*

Il resoconto di *Nature* del 1989 è disponibile in rete [6]. Fra i molti autori, a occuparsi della elaborazione statistica finale furono quelli del British Museum di Londra. Il coordinatore era Michael Tite. I calcoli furono presumibilmente eseguiti da Janet C. Ambers, Sheridan G.E. Bowman e Morven N. Leese. La Ambers e la Leese partecipavano, si suppone, in quanto esperte di statistica. La Bowman è stata direttrice del laboratorio del radiocarbonio (con metodo tradizionale) del British Museum, oggi dismesso.

### §1. I laboratori e i tessuti

Alla datazione del 1988 parteciparono tre laboratori: Università dell'Arizona, Tucson; Università di Oxford; Politecnico (Eidgenössische Technische Hochschule) di Zurigo. Tutti usavano il metodo AMS (Spettrometria di Massa con Acceleratore).

Oltre alla Sindone di Torino, indicata come campione n. 1 nel resoconto, furono datati altri tre tessuti di lino con queste provenienze:

- n. 2 - Qasr Ibrim (Bassa Nubia, Egitto)
- n. 3 - antica Tebe (oggi Luxor, Egitto)
- n. 4 - Saint-Maximin-la-Sainte-Baume (Provenza, Francia)

Di questi era già nota o stimabile l'età. Il n. 3 era di epoca romana, gli altri due di epoca medievale. Per questi tre tessuti, i risultati furono in ottimo accordo con le precedenti stime. Le operazioni furono coordinate dal British Museum di Londra dove fu anche condotta l'elaborazione statistica finale.

## §2. Le procedure per la datazione col metodo AMS

Vediamo in sintesi quali sono (o quali erano all'epoca) le procedure seguite. Quando un campione da datare arriva in laboratorio, subisce un trattamento in diverse fasi preliminari prima della effettiva misurazione del C14.

(1) Pretrattamento. Il campione viene sottoposto a pulitura con uno o più di diversi metodi possibili, di natura fisica e/o chimica.

(2) Combustione per produrre CO<sub>2</sub>.

(3) Grafitizzazione per ottenere carbonio puro in forma di grafite.

La grafite viene compressa in pasticche che vengono collocate lungo il bordo di un tamburo (una ruota) alla maniera dei proiettili sul tamburo di una rivoltella. Il tamburo viene introdotto nello spettrometro. All'epoca i laboratori usavano tamburi da dieci o venti pasticche. Oggi si usano tamburi più grandi, per esempio con cinquanta pasticche o anche fino a cento o più.

(4) Ionizzazione. Le pasticche vengono bombardate con atomi di cesio e si formano atomi di carbonio ionizzati.

A questo punto si passa alla misurazione. Gli ioni di carbonio vengono accelerati su un percorso al termine del quale i vari isotopi si trovano spazialmente separati. Gli atomi di C14 vengono contati uno per uno. Per quelli di C13 o C12 si misura la corrente elettrica portata dagli ioni.

Terminata una misurazione su una pasticca (che può durare all'incirca un minuto), immediatamente si ripete la misurazione sulla stessa pasticca per diverse volte. Al termine il tamburo ruota di un tratto e porta la successiva pasticca in posizione di lancio degli ioni. Si procede per tutte le pasticche del tamburo. Al termine di un giro di tamburo, si prosegue con un nuovo giro, sempre con le stesse pasticche, e si ripetono le misurazioni. Si procede poi ancora con altri giri completi di tamburo. Alla fine di tutto il processo, ciascuna pasticca è stata misurata molte volte, per esempio cinquanta o sessanta volte (come sarebbe con dieci volte a ogni giro per cinque o sei giri).

Sul tamburo vengono sempre montate, assieme alle pasticche di campioni da datare, anche pasticche di "standard" con contenuto noto di C14 e pasticche di "bianco" prive di C14. Possono esserci anche pasticche di altri campioni di età nota da usare come ulteriori controlli.

Per ogni pasticca di campione da datare, si misura il rapporto fra C14 e C13 (o fra C14 e C12) e si fa il confronto con l'analogo rapporto per le pasticche degli standard. Dal confronto si deduce la proporzione di C14, quindi l'età in anni radiocarbonici. Le pasticche dei bianchi servono per misurare la presenza di un fondo (background) dovuto a C14 spurio, una presenza inquinante che andrà sottratta.

Arizona e Oxford misuravano il rapporto C14/C13, mentre Zurigo misurava il rapporto C14/C12.

Per un lavoro accurato, occorre inoltre misurare il rapporto C13/C12 per stimare la correzione di frazionamento isotopico da apportare. Arizona misurava questo rapporto separatamente con un normale spettrometro di massa (senza acceleratore). Il laboratorio di Oxford all'epoca non era ancora attrezzato per questa misura e inviava i campioni a un altro laboratorio inglese (Harwell). Zurigo misurava il C13/C12 quasi simultaneamente al C14/C12 sulla stessa macchina con acceleratore.

Le macchine attuali più avanzate misurano sequenzialmente tutti i rapporti, sia C14/C13 sia C14/C12, oltre a C13/C12.

### §3. Presentazione del risultato

Per una singola datazione, il risultato viene presentato come un intervallo per l'età espressa in anni BP (anni radiocarbonici convenzionali). L'intervallo ha questa forma:

$$x \pm \sigma$$

Il valore centrale  $x$  dell'intervallo è la media delle diverse misurazioni effettuate sullo stesso campione (o sottocampione) ed è una stima dell'età. Il numero sigma dopo il più/meno è la semiampiezza dell'intervallo ed esprime l'incertezza della misurazione. Parliamo qui di “sigma dichiarata” per sottolineare che è la sigma stimata dal laboratorio e non è detto che coincida con la sigma effettiva. Questa sigma dichiarata è (o sarebbe in teoria) la deviazione standard della distribuzione degli errori. Se la distribuzione è quella che in matematica si chiama “normale” (o “gaussiana”), allora l'intervallo va così interpretato. C'è una probabilità del 68% che la vera età (sempre in anni BP) sia compresa entro  $(x \pm \sigma)$  e c'è una probabilità del 95% che sia compresa entro  $(x \pm 2\sigma)$ . Per esempio, con l'intervallo  $900 \pm 50$  c'è una probabilità del 68% che la vera età sia compresa fra 850 e 950, mentre c'è una probabilità del 95% che sia compresa fra 800 e 1000. Quindi quelli solitamente mostrati in prima istanza (come per il resoconto di Nature nella sua prima tabella) sono intervalli di fiducia al 68%. In un secondo tempo si applica la correzione di calibrazione e si passa dalle età in anni BP alle date di calendario, per le quali si fornisce solitamente un intervallo al 95 per cento di fiducia.

Nel presente articolo, dove non altrimenti specificato, le età saranno sempre espresse in anni BP (anni radiocarbonici convenzionali non calibrati).

### §4. La sigma dichiarata

Ho tradotto come “*sigma dichiarata*” quello che in inglese è chiamato “*quoted error*”. È importante notare che la sigma presentata coi risultati tende (o almeno tendeva all'epoca) a sottostimare la reale entità degli errori. Infatti è una sigma dichiarata dal laboratorio che l'ha stimata solo in base alla variabilità riscontrata nell'ultima fase della procedura di datazione.

Abbiamo visto che ciascuna pasticca viene misurata molte volte. Dalla dispersione, o variabilità, delle molte misurazioni il laboratorio stima il valore di sigma da dichiarare. Per il calcolo della sigma dichiarata vengono anche aggiunti (in somma quadratica) altri termini di minore entità per altre incertezze, in particolare per l'incertezza nella misurazione del delta-C13.

Però il laboratorio non può quantificare, per ciascuna pasticca, gli effetti di altri eventuali fattori intervenuti nelle complesse fasi precedenti alla misurazione. Per esempio, può essere che l'inquinamento presente in origine sul campione non sia stato del tutto eliminato con le procedure di pulizia o per converso che sia stato aggiunto nuovo inquinante presente in laboratorio, oppure che si siano verificati fenomeni di frazionamento isotopico (arricchimento selettivo di un isotopo rispetto a un altro). Questi aspetti non possono essere quantificati dal laboratorio, per lo meno non durante la misurazione su una singola pasticca in una singola seduta.

Inoltre anche durante la misurazione possono entrare in gioco fattori non controllabili, per cui le misurazioni di due sottocampioni dello stesso reperto in due diverse sedute, o perfino di due pasticche durante la stessa seduta, possono avere oscillazioni non previste dalle sigma dichiarate.

Per capire se un valore di sigma è realistico oppure è sottostimato (o eventualmente sovrastimato) bisognerebbe ripetere diverse volte la datazione su porzioni dello stesso reperto, cioè ripeterla *in tutte le fasi fin dall'inizio*, compreso il pretrattamento, e in diverse sedute, poi osservare la variabilità nei risultati. Ma questo tradizionalmente non veniva fatto. Inoltre, per tener conto di possibili errori sistematici, il laboratorio dovrebbe regolarmente misurare, con le stesse procedure, reperti per i quali l'età è nota esattamente e, all'occasione, dovrebbe effettuare confronti su uguali reperti con altri laboratori. Oggi queste precauzioni vengono prese dai migliori laboratori e si sono

fatti molti progressi ai fini del controllo di qualità.

Negli anni '80 il controllo di qualità non era ancora perfezionato. Quando, a partire proprio da quegli anni, furono avviati programmi di intercomparazione fra diversi laboratori ai quali venivano fornite diverse porzioni degli stessi reperti, si vide che non di rado le sigma dichiarate dai laboratori erano sottostimate. Si provò a introdurre un fattore moltiplicativo, o moltiplicatore (*multiplier*), per esempio di 1,5 o 2, per adeguare le sigma in base all'esperienza di ciascun laboratorio. Oggi il controllo di qualità dei laboratori è di molto migliorato e le sigma dichiarate dovrebbero essere più realistiche, in quanto i laboratori cercano di tener conto di tutte le possibili fonti di errore..

## §5. Dispersione dei risultati

Dicevamo che per conoscere la *effettiva* incertezza di una datazione, non quella stimata dal laboratorio in base solo alla variabilità delle ripetute misurazioni su una stessa pasticca durante una stessa seduta, occorre ripetere molte volte la datazione di un reperto, cioè di diverse porzioni dello stesso reperto, passando ogni volta attraverso tutte le fasi a partire dal pretrattamento. Si avrebbe così una stima effettiva della variabilità fra diverse ripetizioni. Inoltre bisognerebbe anche operare su campioni di età nota per tener conto di effetti sistematici.

Essendo in possesso di diversi risultati di datazioni distinte (ripetute in tutte le fasi) sullo stesso reperto, dal confronto fra i risultati si può stimare l'effettiva entità degli errori. Lo si fa calcolando una misura della *dispersione* (“*scatter*”) dei risultati. Come misura della dispersione si usa appunto la deviazione standard, grandezza che è di solito indicata con la lettera sigma. Per evitare confusione, useremo di solito la parola esplicita, “dispersione”, e dovendo usare un simbolo la indicheremo con sigma con zero ( $\sigma_0$ ). Occorrerà fare attenzione a non confondere nelle formule le  $\sigma_i$  che sono le sigma dichiarate (*quoted error*) per ciascun risultato, con la  $\sigma_0$  che è la dispersione (*scatter*) di un insieme di risultati.

Se il numero di datazioni distinte sullo stesso reperto fosse abbastanza grande, si potrebbe avere, calcolando la dispersione, una buona stima della *effettiva* deviazione standard della distribuzione degli errori. Per la Sindone abbiamo dodici valori, che non è un numero grande ma è sufficiente per cominciare ad avere una stima indicativa della effettiva deviazione standard.

Avremo occasione di calcolare la dispersione anche per gruppi più piccoli di dati e allora la stima sarà molto incerta ma potrà servire per fare confronti.

Quindi la sigma dichiarata (*quoted error*) e la dispersione (*scatter*) sono due stime della stessa grandezza, cioè della deviazione standard della distribuzione degli errori per ogni singola datazione. Confrontando, per un insieme di datazioni, i valori delle sigma dichiarate per le singole datazioni con il valore della dispersione calcolata sull'insieme, si ha una indicazione in merito all'adeguatezza delle sigma dichiarate. Il confronto è tanto più attendibile quanto più alto è il numero di datazioni ripetute sullo stesso reperto. Se le singole sigma dichiarate sono mediamente all'incirca uguali alla dispersione calcolata sull'insieme, allora si può dire che non sono né sottostimate né sovrastimate.

Aggiungo che il confronto fra diverse datazioni dello stesso reperto, condotte nello stesso laboratorio, fornisce indicazioni sulla precisione ma non sull'accuratezza. Se il laboratorio fosse affetto da un errore sistematico, ciò non verrebbe rilevato se non con datazioni di controllo su reperti di età precedentemente nota o con confronti fra diversi laboratori. (I risultati sono precisi se sono vicini fra loro, mentre sono accurati se sono vicini alla vera età.)

## §6. Calcolo della dispersione

Si calcola come stima della deviazione standard degli  $n$  risultati  $x_i$ . La varianza, che è il quadrato della deviazione standard, si stima con questa formula:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad (1)$$

Nella sommatoria al numeratore ci sono gli scarti dalla media, che è la media semplice (non ponderata):

$$\bar{x} = \frac{\sum x_i}{n} \quad (2)$$

Per ottenere la deviazione standard si estrae la radice quadrata della varianza.

Anche una modesta calcolatrice tascabile ha una funzione per calcolare la deviazione standard, anzi ne ha due perché quella che usiamo è una formula per la varianza corretta per un fattore  $(n/(n-1))$ . È la formula che, fra le due, fornisce un valore leggermente più grande.

## §7. Risultati

La Tabella 1 riporta i risultati per Torino come pubblicati su *Nature*. I risultati per gli altri tessuti si possono vedere in [6].

Le età sono espresse in anni BP (anni radiocarbonici convenzionali, non calibrati)

Tabella 1 – Risultati per Torino

	Arizona	Oxford	Zurigo
	591±30	795±65	733±61
	690±35	730±45	722±56
	606±41	745±55	635±57
	701±33		639±45
			679±51
media	646±31	750±30	676±24

Il primo gruppo di righe, preso dalla Tabella 1 di *Nature*, riporta i risultati per ciascuna singola datazione. Arizona aveva diviso il suo campione in quattro sottocampioni, Oxford in tre, Zurigo in cinque.

L'ultima riga, presa dalla Tabella 2 di *Nature*, riporta le medie dei risultati per ciascun laboratorio. Ne parleremo più avanti.

## §8. Distribuzione dei singoli risultati per Torino

Calcolando la dispersione stimata dai dodici risultati per Torino, si trova 61,2, che è un buon valore, cioè non è troppo grande, per quel che ci si poteva aspettare all'epoca.

Vediamo come sono distribuiti i dodici risultati. Nella Tabella 2 sono mostrati gli scarti, cioè le distanze dalla media (semplice) dei dodici valori per le età di Torino. A fianco di ogni valore è indicato il laboratorio. Le linee tratteggiate separano i valori alle distanze dalla media di una o due deviazioni standard supposte uguali a 61. I numeri nell'ultima colonna sono i rapporti fra gli scarti della prima colonna e il valore della deviazione standard, cioè sono i valori della variabile  $z$  (normale standardizzata).

Tabella 2 - Scarti dalla media dei 12 risultati

scarti	z	
-----		-122
-97,8 A	-1,60	
-82,8 A	-1,35	
-----		-61
-53,8 Z	-0,88	
-49,8 Z	-0,81	
-9,8 Z	-0,16	
1,2 A	0,02	
12,2 A	0,20	
33,2 Z	0,54	
41,2 O	0,67	
44,2 Z	0,72	
56,2 O	0,92	
-----		61
106,2 O	1,73	
-----		122

Supponendo una sigma effettiva di 61 anni (come trovato per la dispersione), si vede che la situazione è come ci si aspetta. Entro la distanza di una deviazione standard, in teoria ci dovrebbero essere circa il 68% di 12, cioè 8 valori. Ce ne sono 9. Nessun valore cade al di là di una distanza di due deviazioni standard. Siamo comodamente nell'attesa statistica. I dati sono ben distribuiti e non ci sono risultati anomali.

Se però si guarda a come sono distribuiti i risultati di ciascun laboratorio, si vede che i tre risultati di Oxford si collocano all'ultimo, penultimo e quartultimo posto, dalla parte delle età più antiche. Questo farebbe pensare che Oxford soffrisse di un errore sistematico che tendeva a produrre età più antiche. Però per gli altri tre tessuti non c'è un simile effetto, o non così marcato. Vedremo più avanti che, tenendo conto dei valori di delta-C13, si potrebbe sospettare un effetto sistematico per Oxford, ma è troppo lieve per poter decidere di scartare i valori di Oxford. Se si potesse farlo, la dispersione si ridurrebbe a un eccellente valore di 50,6. Ma non si possono eliminare i valori di Oxford perché potrebbero essere dovuti solo a un effetto casuale, benché poco probabile. Oxford aveva tre soli valori e poteva succedere che cadessero verso le età più vecchie. Se uno solo fosse risultato più giovane, la differenza sarebbe sparita. Un altro motivo per non potere eliminare Oxford è che, con soli tre laboratori, non possiamo escludere che, viceversa, siano gli altri due ad avere un effetto sistematico verso le età più giovani. Comunque la cosa ha poca importanza perché eliminando Oxford la media (semplice) di tutti i risultati si sposterebbe di soli 22 anni verso le età più giovani.

Non è il caso di preoccuparsi per la differenza fra Oxford e gli altri laboratori perché all'epoca, negli anni '80, quando i laboratori con AMS erano da poco entrati nell'uso (ma in realtà anche per i laboratori con metodo convenzionale), non c'era da meravigliarsi se confrontando i risultati di diversi laboratori si trovavano differenze di cento o anche più anni BP.

A me sembra che, più che parlare di sfortuna per il confronto fra laboratori per Torino, bisognerebbe parlare di fortuna per gli altri tre tessuti, per i quali le medie di ciascun laboratorio sono risultate molto vicine fra loro. Una simile concordanza non era certo la regola, all'epoca, se si facevano

confronti fra diversi laboratori per la datazione dello stesso reperto.

In realtà i laboratori quella volta poterono essere ben contenti del risultato complessivo. Potevano temere di peggio. Qualche anno prima, nel 1983-1984, gli stessi tre laboratori, assieme ad altri tre, avevano effettuato datazioni sugli stessi tessuti per vedere se avevano risultati concordanti e quindi per candidarsi per la datazione della Sindone. I risultati furono notevolmente peggiori di quelli poi ottenuti nel 1988 (li vedremo nella Parte III), ma anche così i laboratori ritennero che la concordanza fosse soddisfacente.

## §9. Il delta-C13

Misurando il rapporto C13/C12, e confrontandolo con uno standard, si calcola il delta-C13 ( $\delta^{13}\text{C}$ ) che serve per la correzione di frazionamento isotopico. Per qualche nozione sul frazionamento isotopico in relazione al radiocarbonio, vedere per esempio [17].

Per le piante più comuni, e per il lino, il delta-C13 è di solito attorno al valore di -25 per mille. Significa che il rapporto C13/C12 è inferiore del 25 per mille rispetto allo standard. I calcoli dei laboratori del radiocarbonio sono normalizzati a -25 (s'intende sempre per mille). Se sul campione da datare si misura un delta di -25, non si apporta una correzione. Se si misura un delta diverso, si apporta una correzione che è di 16-17 anni per ogni punto per mille di differenza. Se il delta misurato è inferiore a -25, per esempio -27, la correzione apportata va nel senso di una età inferiore, ovvero di una data più recente.

La Tabella 3 riporta i valori di delta-C13 misurati dai tre laboratori per i quattro tessuti. I dati sono presi dalla Tabella 1 di *Nature*.

Tabella 3 – Valori di delta-C13

	Torino	Nubia	Tebe	Provenza
Arizona	-25	-23	-23,6	-25
Oxford	-27	(-27)	-27	(-27)
Zurigo	-25,1	-23,6	-22	-25,5

Per Oxford i valori di delta per Nubia e per Provenza non sono stati misurati e sono solo stati supposti uguali a -27. Abbiamo messo in parentesi i due valori. È strano che Oxford fosse tanto in economia da far misurare il delta-C13 (nel laboratorio di Harwell) per soli due tessuti, cioè Torino e Tebe. Per entrambi trovò -27 per mille e suppose lo stesso valore anche per Nubia e Provenza.

I risultati (le età) forniti per le datazioni incorporano già le correzioni di frazionamento isotopico.

Quindi Oxford con -27 aveva trovato, prima della correzione, una età ancora più vecchia di quella pubblicata, poi con la correzione ha portato avanti la data di circa 33 anni.

Mentre Arizona e Zurigo trovano, per ciascun tessuto, valori di delta simili o non troppo lontani, per Oxford il valore di -27, misurato o supposto, è piuttosto lontano dagli altri due laboratori.

Le differenze non sono di poco conto. Lo scarto più grosso nei valori di delta-C13, quello di cinque per mille fra Oxford e Zurigo per Tebe, comporta una differenza di 83 anni nella data.

Per Torino, la media del delta dei laboratori di Arizona e Zurigo è -25,05, con una differenza rispetto a Oxford di 1,95, a cui corrisponde una differenza di 32 anni nella correzione per le età. Quindi per Oxford, che già ha trovato età più vecchie per Torino, un delta di -27 comporta che

prima della correzione le sue età erano ancora più vecchie. Però anche per gli altri tessuti la correzione, per Oxford, porta a una età più giovane, e in due casi, se si fa il confronto con il delta degli altri laboratori, la differenza è notevole. Per Nubia, gli altri due laboratori hanno -23 e -23,6, quindi con una differenza rispetto a Oxford in media di 3,7 per mille. Per Tebe, -22 e -23,6, la differenza rispetto al -27 di Oxford è ancora maggiore, in media di 4,2 per mille

Naturalmente il reale valore di delta-C13, per ciascun tessuto, dovrebbe essere lo stesso per tutti i laboratori. Non sappiamo se sia stato Oxford (cioè Harwell) a trovare un valore troppo piccolo, o se l'errore sia negli altri due laboratori, ma gli altri due laboratori trovano valori vicini fra loro e sembra lecito, in via di ipotesi, supporre che i risultati anomali siano quelli di Oxford. Di certo, per Nubia e Provenza, dove Oxford non ha misurato il delta-C13, è logico assumere, invece del -27 assunto da Oxford, una media fra i valori misurati da Arizona e Zurigo, anzi ci si chiede perché questo non sia stato fatto durante l'elaborazione finale dei dati al British Museum. Possiamo ricalcolare i risultati per le età di Oxford facendo l'ipotesi che Oxford avesse trovato, per ciascun tessuto, un delta-C13 uguale alla media dei valori trovati dagli altri due laboratori. Con i dati ricalcolati, facciamo il confronto fra la media delle età di Oxford e la media delle età degli altri due laboratori. Risulta il quadro come in Tabella 4. I calcoli sono stati fatti usando l'età media di ciascun laboratorio, cioè la media ponderata quale pubblicata su *Nature*.

Tabella 4

	Oxford	Altri	diff.	Oxford	Altri	diff.
Torino	750	661	+89	782,4	661	+121,4
Nubia	940	934	+6	1001,4	934	+67,4
Tebe	1980	1967,5	+12,5	2049,7	1967,5	+82,2
Provenza	755	703,5	+51,5	784,1	703,5	+80,6

Nelle prime tre colonne, i dati sono quelli pubblicati su *Nature* e non sono stati ricalcolati. Nella prima colonna ci sono le età di Oxford. Nella seconda colonna la media delle età di Arizona e Zurigo. Nella terza colonna c'è la differenza, con il segno più che indica che l'età di Oxford è più vecchia. Si vede che Oxford ha sempre una età più vecchia, per tutti i quattro tessuti, ma le differenze possono essere piccole.

Passiamo al secondo gruppo di tre colonne, dove i dati sono ricalcolati usando per Oxford un delta-C13 uguale alla media degli altri due laboratori anziché il -27. Nell'ordine ci sono le età di Oxford dopo il ricalcolo, la media degli altri due laboratori (come nella seconda colonna) e la differenza. Ora si vede che per Torino la differenza diventa molto forte, con l'età di Oxford che è oltre 120 anni più vecchia. Però anche per gli altri tre tessuti c'è una differenza nel senso di età che per Oxford sono sempre più vecchie.

Le differenze non sono importanti ma si potrebbe supporre la presenza di un lieve errore sistematico da parte di Oxford che tendeva a produrre valori maggiori per l'età (oppure, viceversa, età più giovani per gli altri due laboratori).

Va anche aggiunto che le differenze sono un po' inferiori se nel ricalcolo si usano le medie semplici anziché ponderate.

L'incertezza nella misurazione del delta-C13 va incorporata nella sigma dichiarata, dando in genere solo un piccolo contributo. Va aggiunto (in somma quadratica) un termine corrispondente all'incertezza della correzione da applicare per il frazionamento isotopico. In questo modo la sigma dichiarata risulta leggermente aumentata. Oxford e Zurigo dovrebbero aver tenuto conto di questa incertezza nel calcolo delle sigma dichiarate da loro fornite. Va notato che invece Arizona, secondo il resoconto di *Nature*, ha fornito le sue sigma dichiarate senza ancora tener conto dell'incertezza nel delta-C13. Tale incertezza è stata introdotta nei calcoli successivi, non si sa esattamente in che modo. È strano che i laboratori non si fossero accordati per adottare una procedura uniforme.

Nessuno dei tre laboratori indica quale fosse l'entità del termine aggiuntivo usato per l'incertezza del



delta-C13.

## §10. Confronto fra le dispersioni per i quattro tessuti

Nella Tabella 5 abbiamo cumulato, per ciascun tessuto, tutti i risultati delle singole datazioni per i tre laboratori, che sono in numero fra 11 e 13. Abbiamo messo a confronto i quattro tessuti.

Tabella 5

	n	dispers	dichiar	rapporto	estremi
1) Torino	12	61,22	47,83	1,28	204
2) Nubia	13	55,71	49,92	1,12	207
3) Tebe	11	76,60	50,09	1,53	299
4) Provenza	13	65,64	52	1,26	223

Nella prima colonna c'è il numero totale di risultati per ciascun tessuto. Nella seconda colonna la dispersione. Nella terza colonna la media (semplice) delle sigma dichiarate. Nella quarta colonna il rapporto fra dispersione e media delle sigma dichiarate. Nell'ultima colonna la differenza fra gli estremi, cioè fra la età più vecchia e la più giovane.

Si vede che la dispersione per Torino non è la peggiore. Solo uno degli altri tre tessuti, Nubia, ha una dispersione più piccola.

La differenza fra gli estremi per Torino è la più piccola.

La dispersione è in tutti i casi più grande della media delle sigma dichiarate, quindi le sigma dichiarate tendono a essere sottostimate rispetto ai reali errori, anche se di poco. Un fattore moltiplicativo compreso attorno a 1,2-1,5 sarebbe modesto e non tale da sorprendere.

Per Torino, la differenza fra media delle sigma dichiarate e dispersione sarebbe un po' minore, in pratica molto piccola, se non ci fosse la riduzione delle sigma dichiarate per Arizona per i motivi che vedremo al §20.

## §11. Altri confronti

Vediamo alcune tabelle che mostrano risultati distinti per tessuto e per laboratorio.

Tabella 6 - Media delle sigma dichiarate

	Torino	Nubia	Tebe	Provenza	media
Arizona	34,75	45,8	45,6	51,4	44,39
Oxford	55	51,67	58,33	45	52,50
Zurigo	54	53	49,33	56,8	53,28

La Tabella 6 mostra, per ciascun laboratorio e ciascun tessuto, la media delle sigma dichiarate. Nell'ultima colonna c'è la media dei valori sulla stessa riga. Si vede che il dato per Arizona/Torino è nettamente più basso, e capiremo il perché al §20.

Tabella 7 - Dispersione

	Torino	Nubia	Tebe	Provenza	media
Arizona	56,52	68,82	109,83	84,46	79,91
Oxford	33,87	35	17,56	44,81	32,81
Zurigo	45,46	60,59	50,21	50,35	51,65

La Tabella 7 mostra la dispersione fra i risultati per ciascun laboratorio e ciascun tessuto. Si vede che i dati di Arizona sono sempre quelli con maggiore dispersione, cioè con variazioni maggiori fra i risultati, in particolare per Tebe e Provenza.

Tabella 8 - Differenza fra estremi

	Torino	Nubia	Tebe	Provenza	media
Arizona	110	167	299	223	199,75
Oxford	65	65	35	80	61,25
Zurigo	98	146	98	114	114

La Tabella 8 mostra la differenza fra il valore massimo e il valore minimo delle età per ciascun laboratorio e ciascun tessuto. Naturalmente si conferma il dato per la dispersione. Arizona ha i valori maggiori per la differenza fra estremi, in particolare per Tebe e Provenza. Oxford ha i valori minori. Va considerato che i dati di Oxford sono soltanto tre per ciascun tessuto, quindi aggiungendo uno o due dati, la differenza fra estremi potrebbe aumentare.

Tabella 9 - Test chi-quadrato

	Torino	Nubia	Tebe	Provenza
Arizona	0,034	0,06	0,0002	0,002
Oxford	0,71	0,66	0,92	0,33
Zurigo	0,60	0,35	0,35	0,56

Come diremo al §15, il test chi-quadrato serve per misurare la omogeneità dei risultati, ovvero la compatibilità fra la variabilità dei risultati e le sigma dichiarate. Sono mostrati i valori di  $p$ , un numero (livello di significatività del test) che può variare fra zero e uno e che è tanto più piccolo (vicino a zero) quanto minore è la compatibilità.

In generale, se il valore di  $p$  è centrale (attorno a 0,50), la variabilità è quella che ci si aspetta con quei valori di sigma dichiarata.

Se il valore di  $p$  sale verso l'unità, la variabilità è più piccola di quanto ci si aspetta, come dire che i vari risultati sono troppo omogenei. Significa che le sigma dichiarate sono sovrastimate.

Se il valore di  $p$  scende verso lo zero, allora la variabilità è maggiore rispetto a quanto ci si attende con quei valori di sigma dichiarata. Cioè i risultati appaiono non omogenei. Quindi le sigma dichiarate sono sottostimate. Per convenzione, si dice che se  $p$  scende sotto 0,05 la disomogeneità è troppo forte (per quei valori di sigma dichiarata) e il test di omogeneità del chi-quadrato non è superato. (Tenicamente, nel test di verifica di una ipotesi,  $p$  è la probabilità di errore di I Tipo.)

Anche per una  $p$  di poco superiore a 0,05, per esempio inferiore a 0,10, è bene sospettare che le sigma siano sottostimate.

Come si vede dalla tabella 9, per Oxford e Zurigo il test è comodamente superato e non c'è evidenza che le sigma dichiarate siano sottostimate. Invece per Arizona, in particolare con Tebe e Provenza, la variabilità dei risultati è maggiore di quanto ci si può aspettare dai valori delle sigma dichiarate, ovvero le sigma dichiarate appaiono sottostimate.

## §12. Era necessario combinare i risultati?

Se uno o più laboratori effettuano diverse datazioni su porzioni di uno stesso reperto, come nel caso della Sindone e dei tre tessuti di controllo, e se si vuole presentare un risultato espresso formalmente come un singolo intervallo di fiducia che tenga conto di tutte le misurazioni, occorre combinare i risultati.

Descriveremo i metodi usati da *Nature* per combinare i risultati. Occorre però avvisare che per questi metodi, o in generale per altri metodi che si potrebbero usare, le conclusioni vanno prese non alla lettera ma come generiche indicazioni alle quali è opportuno anteporre un “circa”. Infatti tali metodi sarebbero esattamente validi solo assumendo stringenti condizioni che in genere non sono esattamente verificate, o almeno non erano verificate all'epoca. Occorre che la distribuzione di probabilità per i risultati di ciascuna datazione sia quella che in matematica si chiama “distribuzione normale” (o “gaussiana”). Occorre che per tutte le datazioni la distribuzione sia centrata sullo stesso valore, la “vera” età del reperto. Occorre che l'incertezza o sigma di ciascuna datazione sia nota. Non è realistico pensare che tutte queste condizioni siano sempre verificate nella pratica. I singoli risultati possono avere una distribuzione irregolare che non si conforma esattamente a una distribuzione normale. Possono esserci piccoli errori sistematici tali che non tutte le distribuzioni sono centrate sullo stesso valore. Le sigma delle distribuzioni non sono note con esattezza. Quindi c'è un rischio concreto, che volendo calcolare un unico intervallo di fiducia, si ottenga come risultato un intervallo per il quale sarebbe ottimistico presumere di conoscere esattamente gli estremi e il livello di fiducia. In particolare, l'intervallo di fiducia potrebbe risultare meno ampio di quanto dovuto. Infatti i metodi per combinare i risultati comportano tutti una riduzione dell'ampiezza dell'intervallo di fiducia, rispetto agli intervalli dichiarati per le singole datazioni (che già essi possono essere sottostimati), e la riduzione potrebbe non essere giustificata se le condizioni non sono tutte verificate.

In certe situazioni, se si prova a combinare gli stessi dati con diversi metodi, tutti in teoria legittimi, si possono trovare risultati diversi, come riprova del fatto che occorre usare molta elasticità nel valutare le conclusioni.

A mio parere, il modo più opportuno per presentare i risultati per la Sindone sarebbe stato di fornire distintamente tutti i dodici risultati delle singole datazioni in anni BP, assieme alle rispettive date calibrate, senza tentare di combinarli in un unico intervallo di fiducia, poi lasciare l'interpretazione al buon senso.

I dodici risultati in anni BP per la Sindone corrispondono a date calibrate che vanno circa dal 1250 al 1350. Quindi l'esito della datazione si può esprimere dicendo che probabilmente il telo risale a un periodo fra la seconda metà del Duecento e la prima metà del Trecento, o, volendo essere più sicuri, fra l'inizio del Duecento e la fine del Trecento. Questo è più che sufficiente per concludere che la Sindone risale al medioevo e non all'epoca di Cristo. Una precisione maggiore non si può ottenere con i dati a disposizione, per esempio non si può dire che la Sindone sia del 1280 o del 1340.

È inutile aggiungere che il risultato è in accordo con le notizie storiche sulla comparsa della Sindone.

Ripetiamo che l'esito fu molto soddisfacente, per quel che si poteva ottenere allora, per il complesso delle datazioni dei quattro tessuti. Gli specialisti del radiocarbonio, in particolare quelli del metodo AMS, potevano temere di peggio e devono avere avuto un sospiro di sollievo quando hanno visto i risultati. Qualcuno del mestiere parlò di un “trionfo” per il metodo AMS.

Perché ci si può fidare che il risultato sia soddisfacente? L'esperienza delle datazioni di quegli anni, per quei casi in cui era possibile verificare il risultato, mostra che ci potevano essere errori anche di più di cento anni, talvolta di duecento anni, ma non di mille o più anni. Un risultato anomalo e fortemente sbagliato poteva aversi in casi particolari in seguito a qualche incidente, ma non ci si aspetta che si verifichi per dodici volte in tre laboratori. Non si può ipotizzare un errore sistematico di molti secoli in tutti i laboratori perché i tre tessuti di controllo, datati nelle stesse condizioni, hanno dato tutti risultati in conformità con le anteriori conoscenze sulle età dei reperti. Quindi l'esito della datazione della Sindone al basso medioevo è pacifico e l'elenco dei dodici singoli risultati è sufficiente per dimostrarlo.

Ma gli autori di *Nature* hanno voluto combinare i risultati arrivando a un unico intervallo di fiducia. Senza discutere oltre se la scelta fosse opportuna, ci dedicheremo a illustrare i metodi usati da *Nature* per combinare i risultati.

### §13. Metodi per la combinazione dei risultati di datazioni ripetute

Nel resoconto di *Nature* sono stati usati tre metodi per calcolare la media di diverse datazioni e la relativa incertezza. Li indichiamo per semplicità come Metodo I, II, III.

*Nature* ha scelto di eseguire i calcoli in due fasi, calcolando dapprima la media per le datazioni entro ciascun laboratorio, poi la media delle medie dei tre laboratori. Si poteva anche lavorare sui dati cumulati senza distinguere fra i laboratori, perché fra le assunzioni alla base dei metodi c'è che tutte le datazioni siano stime dello stesso parametro, la vera età del reperto.

Gli autori in partenza intendevano usare quello che qui indichiamo come il Metodo I, ma per applicare questo metodo occorre che sia superato un test per la omogeneità dei risultati da combinare. È stato usato un test del chi-quadrato. Se non viene superato il test, significa che la variabilità dei risultati non è compatibile con le incertezze ( $\sigma$ ) usate per il calcolo della media, ovvero che le  $\sigma$  sono troppo piccole. Gli autori trovano che il test non viene superato in due situazioni, cioè, primo, per le medie entro laboratorio di Arizona per tutti i quattro tessuti e, secondo, per la media delle medie dei tre laboratori per Torino. Ricorrono allora a due ulteriori metodi. Il Metodo II usa un moltiplicatore ottenuto dal chi-quadrato. Il Metodo III si basa sulla dispersione dei risultati. Il Metodo II è stato usato per le medie entro laboratorio dei quattro tessuti di Arizona. Il Metodo III è stato usato per la media delle medie dei tre laboratori per Torino.

Il resoconto di *Nature* non è un modello di chiarezza e non spiega in dettaglio i metodi usati. In realtà per le medie entro laboratorio di Arizona, è una mia interpretazione che sia stato usato il Metodo II. La cosa mi sembra molto probabile ma gli autori non dicono niente in proposito.

Il calcolo delle medie entro laboratorio è forse stato fatto da ciascuno dei laboratori prima di inviare i dati a Londra, ma in ogni caso le tre autrici di Londra avrebbero dovuto controllare i calcoli e decidere i metodi. Quindi direi che la responsabilità della parte statistica è da attribuire alle tre autrici citate.

### §14. Metodo I: Media ponderata usando le sigma dichiarate

Il primo metodo consiste nel calcolare la media ponderata delle varie stime dell'età, usando come pesi gli inversi delle varianze, cioè dei quadrati delle  $\sigma$  dichiarate o stimate. Se i risultati da combinare hanno la forma:

$$x_i \pm \sigma_i$$

dove le  $x_i$  sono le varie stime dell'età e le  $\sigma_i$  sono le rispettive incertezze, la media ponderata delle stime dell'età si calcola così:

$$m = \frac{\sum (x_i / \sigma_i^2)}{\sum (1 / \sigma_i^2)} \quad (3)$$

e l'incertezza o sigma della media ponderata si calcola così:

$$\sigma_m^2 = \frac{1}{\sum (1 / \sigma_i^2)} \quad (4)$$

Quella indicata è la sigma al quadrato, cioè la varianza, poi si farà la radice quadrata per avere la sigma.

Se le sigma dei vari risultati da combinare fossero tutte uguali, allora la media ponderata sarebbe la media semplice e l'incertezza della media sarebbe la sigma comune divisa per la radice di  $n$ , dove  $n$  è il numero dei risultati da combinare.

Con i valori calcolati per la media ponderata (3) e per la sua sigma che è la radice di (4) (e se è superato il test del prossimo §15) si costruisce come risultato finale l'intervallo di fiducia per il complesso delle varie misurazioni da combinare. L'intervallo è centrato sulla media ponderata. L'ampiezza dell'intervallo dipende dalla sigma. La semiampiezza dell'intervallo è pari a una sigma (per il livello di fiducia al 68%) o a due sigma (al 95%).

Gli autori di *Nature* hanno applicato il Metodo I nella prima fase per le medie entro laboratorio di Oxford e Zurigo, usando per le  $x_i$  le singole stime dell'età e per le  $\sigma_i$  le rispettive sigma dichiarate. Poi nella seconda fase lo hanno applicato per la media delle medie dei tre laboratori per Nubia, Tebe e Provenza, usando per le  $x_i$  le medie e per le  $\sigma_i$  le sigma delle medie quali trovate nella prima fase.

## §15. Un test chi-quadrato

Prima di calcolare l'intervallo di fiducia col Metodo I, cioè prima di considerare attendibile il valore per la sigma della media, occorre verificare se viene superato un test di omogeneità che controlla se la variabilità nelle singole stime delle età è compatibile con le sigma presunte (si tratta per la prima fase delle sigma dichiarate, oppure per la seconda fase delle sigma per la media calcolate a partire dalle dichiarate). *Nature* usa come test un calcolo del chi-quadrato. Il chi-quadrato si calcola così:

$$\chi^2 = \sum \frac{(x_i - m)^2}{\sigma_i^2} \quad (5)$$

dove in parentesi c'è la differenza fra ciascuna stima dell'età e la media ponderata.

Al chi-quadrato è associato un numero di gradi di libertà, che nella nostra situazione, almeno nella prima fase, è dato da  $(n - 1)$ , dove  $n$  è il numero di dati da combinare.

Una volta calcolato il chi-quadrato, si consulta una apposita tabella, o più comodamente si usa uno dei programmi disponibili online, e si trova il corrispondente valore di  $p$ , il livello di significatività. Senza entrare qui nella trattazione della statistica della verifica di ipotesi, diciamo che  $p$  (tecnicamente la probabilità di Errore di I Tipo) è la probabilità che le differenze fra le varie stime dell'età siano dovute solo ai casuali errori di misurazione quali espressi nelle rispettive sigma. La  $p$  può assumere valori compresi fra zero e uno. Se  $p$  è piccola, cioè vicina allo zero, è piccola la probabilità che le differenze fra le stime siano solo casuali. Allora si dice che il valore di  $p$  è "significativo" (cioè le differenze fra le varie età hanno una causa, o un significato, e non sono attribuibili ai soli errori casuali di misurazione previsti dai valori di sigma).

Se si ottiene un valore di  $p$  considerato significativo, allora ci si trova di fronte a due possibili alternative. Primo, il parametro da stimare (la "vera" età dei campioni o sottocampioni) non è sempre lo stesso per i vari campioni o sottocampioni esaminati. Secondo, il parametro è sempre lo stesso ma i valori presunti di sigma sono sottostimati, cioè non sono quelli reali e non rendono

conto della variabilità dovuta a errori di misurazione. Nel nostro caso, la prima alternativa è poco verosimile perché si effettua la datazione di frammenti adiacenti dello stesso tessuto. Invece la seconda alternativa è ben possibile perché è noto che all'epoca succedeva spesso che le sigma dichiarate dai laboratori fossero sottostimate, come abbiamo già notato. Quindi, se le sigma sono sottostimate, proseguendo il calcolo fino a trovare un intervallo di fiducia si ottiene un intervallo non sufficientemente ampio.

Per convenzione, si considera di solito “significativa” una  $p$  inferiore a 0,05 o cinque per cento. Se la  $p$  è molto vicina a zero, per esempio 0,001 o uno per mille, allora la differenza è fortemente significativa. Quindi se si trova una  $p$  inferiore a 0,05 non si accetta l'intervallo di fiducia ottenuto con il Metodo I. Per il nostro problema, anche una  $p$  un po' più grande, diciamo fino a 0,10, indurrebbe secondo me a prudenza prima di accettare l'intervallo di fiducia calcolato con il Metodo I. Invece una  $p$  corrispondente ai valori centrali della distribuzione chi-quadrato, diciamo sopra 0,25, non crea problemi.

È importante notare che se si trova un chi-quadrato con  $p$  significativa, ciò non significa che le datazioni siano prive di validità. Significa solo, nella nostra situazione, che i valori della sigma dichiarata sono sottostimati e non rendono conto della variabilità dei risultati. Non sorprende certo che ci siano casi in cui le sigma dichiarate sono sottostimate. Come abbiamo già detto, i laboratori possono non essere in grado di quantificare tutte le cause di errori sperimentali. Questo era vero in particolare all'epoca della datazione. Negli anni successivi il controllo di qualità è migliorato e i laboratori hanno imparato a cercare di fornire valori un po' più realistici per la sigma dichiarata.

Quindi se il test chi-quadrato non viene superato, allora non significa che i risultati siano da scartare, significa solo che usando il Metodo I si troverebbe un intervallo di fiducia non abbastanza ampio.

Come si procede allora? Per prima cosa si cerca di capire se ci sono uno o più risultati anomali, i cosiddetti “*outliers*”, che si distaccano nettamente dagli altri, tanto da poter essere interpretati come eccezionali e dovuti a qualche incidente imprevisto, per esempio un malfunzionamento di una apparecchiatura o un inquinamento durante le fasi preliminari di trattamento del campione. Se è possibile isolare simili risultati anomali, li si elimina e si procede con gli altri. Non è successo nel nostro caso.

Se non ci sono risultati da scartare, si procede cercando un modo di tener conto del fatto che le sigma dichiarate erano sottostimate. Qui descriviamo due metodi che indichiamo come Metodo II e Metodo III

Gli autori di *Nature* hanno incontrato due situazioni in cui il test non era superato, anche se una delle due non viene citata esplicitamente nel resoconto. Si tratta del calcolo della media per ciascuno dei quattro tessuti entro il laboratorio di Arizona e del calcolo della media delle medie dei tre laboratori per Torino.

Va detto che se il test del chi-quadrato è superato, le conclusioni riguardano solo la precisione della misurazione, ma non l'accuratezza. Se per esempio c'è un errore sistematico nelle misurazioni, i risultati delle misurazioni possono essere vicini fra loro (cioè precisi) e il test viene superato, ma i risultati possono essere lontani dal vero valore del parametro da stimare (cioè non accurati). Per verificare l'accuratezza nella datazione di un reperto di età incognita, occorre eseguire, in condizioni quanto più possibile simili, datazioni di altri reperti di età nota. È quanto è stato fatto, con ottimi risultati, con i tre tessuti di controllo datati assieme alla Sindone,.

## §16. Metodo II: Con il moltiplicatore

In questo metodo, riconoscendo che le sigma dichiarate sono sottostimate, si applica un *moltiplicatore* a tutte le sigma dichiarate in modo da portarle a valori realistici e che siano in grado di spiegare la variabilità dei risultati. Viene usato lo stesso moltiplicatore per tutte le sigma, che quindi conservano gli stessi rapporti l'una con l'altra. Questo significa che, anche se nell'insieme le sigma dichiarate sono sottostimate, si ritiene che i loro rapporti restino indicativi della relativa

precisione di ciascuna misurazione.  
Un opportuno moltiplicatore è questo:

$$\sqrt{\frac{\chi^2}{n-1}} \quad (6)$$

Al numeratore sotto radice c'è il chi-quadrato calcolato con la formula (5). Al denominatore c'è il numero di gradi di libertà.

Si applica lo stesso moltiplicatore a tutte le sigma dichiarate. La media ponderata *resta la stessa* come calcolata con la formula (3) per il Metodo I (vedremo che i sindonologi non lo fanno). Invece la sigma della media risulta anche essa moltiplicata per lo stesso fattore. Quindi in pratica si procede come segue. Si calcola la media ponderata con la formula (3) come nel Metodo I. Si calcola la sigma della media con la formula (4) (estraendo la radice) e la si moltiplica per il fattore (6).

In questo modo l'intervallo di fiducia per la media viene allargato in misura tale da tener conto della variabilità fra i singoli risultati.

Ritengo molto probabile che gli autori di *Nature* abbiano usato questo metodo per correggere le sigma delle medie per Arizona, anche se essi non spiegano quello che hanno fatto. Vedremo in seguito un confronto fra i risultati del calcolo col Metodo II e i valori forniti da *Nature*.

### §17. Metodo III: Con la dispersione

In questo metodo si ignorano completamente le sigma dichiarate (o, se siamo nella seconda fase, le sigma delle medie calcolate dalle sigma dichiarate), ritenute non adeguate, e ci si basa sulla dispersione dei vari risultati per le stime dell'età (o delle varie medie). Si calcola la dispersione, ossia la deviazione standard  $\sigma_0$  delle varie stime dell'età (o delle varie medie) con la formula (1) e si procede come se tutte le sigma fossero uguali alla dispersione. Per la media si usa la media semplice:

$$\bar{x} = \frac{\sum x_i}{n} \quad (2)$$

Per l'incertezza della media (errore standard) si usa questa formula:

$$\frac{\sigma_0}{\sqrt{n}} \quad (7)$$

Sarà bene ripetere che la sigma al numeratore è la dispersione dei risultati, cioè è calcolata con la (1) e non è da confondere con le sigma dichiarate che qui non vengono in alcun modo prese in considerazione.

Per costruire l'intervallo di fiducia finale, occorre tener conto che la vera sigma della distribuzione degli errori non è nota ed è solo stata stimata dai pochi valori dei singoli risultati. Allora si usano i valori critici della variabile  $t$  di Student anziché della variabile  $z$  (normale standardizzata). Significa che una volta calcolato l'errore standard con la (7), la semiampiezza dell'intervallo di fiducia al 95% non viene ottenuta moltiplicando l'errore standard per 2 (o più esattamente per 1,96), ma per il valore che si trova sulle tavole di  $t$  per il numero appropriato di gradi di libertà e per l'area di 0.95 a due code. Per riconoscere facilmente quale colonna usare nelle tavole della  $t$ , basta guardare l'ultimo valore in basso, per infiniti gradi di libertà, che deve essere 1,96. Il valore di  $t$  diventa tanto più grande quanto più è piccolo il numero di gradi libertà. Per esempio, sempre per un intervallo di fiducia al 95%, per 2 gradi di libertà si ha  $t = 4,303$ ; per 9 gradi di libertà  $t = 2,262$ . All'aumentare dei gradi di libertà il valore di  $t$  tende a 1,96, cioè al valore di  $z$ .

Gli autori di *Nature* hanno usato questo metodo per la media delle medie dei tre laboratori per Torino, e questa volta lo hanno detto.

Passiamo a vedere in concreto come sono stati effettuati i calcoli nelle varie situazioni.

### §18. Calcolo delle medie entro laboratorio per Oxford e Zurigo

Qui è stato usato il Metodo I. Era appropriato farlo perché, come si vede dalla Tabella 9, per Oxford e Zurigo e per tutti i tessuti il test del chi-quadrato è ampiamente superato. Nelle due Tabelle 10 e 11 sono messi a confronto i risultati quali forniti da *Nature* (nella sua Tabella 2) con quelli che ho calcolato col Metodo I. Nei miei risultati lascio due decimali. Nel confronto va considerato che naturalmente nei passaggi intermedi dei miei calcoli ho mantenuto abbondanza di decimali, arrotondando a due decimali solo alla fine, però partivo dai numeri forniti nella Tabella 1 di *Nature* dove i valori sono arrotondati all'unità o, per Oxford, alle cinque unità.

Tabella 10 - Medie di Oxford

	Nature	Calcolato
Torino	750 ± 30	749,17 ± 30,70
Nubia	940 ± 30	937,88 ± 29,42
Tebe	1980 ± 35	1977,05 ± 32,71
Provenza	755 ± 30	755,76 ± 25,66

Ci sono differenze di poche unità chiaramente imputabili al fatto che Oxford arrotonda sempre di cinque in cinque unità, sia nella Tabella 1 sia nella Tabella 2 di *Nature*. Però per Provenza c'è una differenza notevole di più di quattro unità per l'incertezza della media. Per inciso, ci si può chiedere perché non sia stata adottata la stessa regola per gli arrotondamenti per i tre laboratori. Il fatto che i dati di Oxford sono arrotondati in modo diverso rispetto agli altri due laboratori, fa pensare che in effetti a Londra abbiano preso i dati inviati da ciascun laboratorio senza intervenire.

Tabella 11 - Medie di Zurigo

	Nature	Calcolato
Torino	676 ± 24	676,14 ± 23,74
Nubia	941 ± 23	940,60 ± 23,16
Tebe	1940 ± 30	1939,81 ± 28,47
Provenza	685 ± 34	685,16 ± 24,63

Zurigo arrotonda all'unità, sia nella Tabella 1 che nella Tabella 2 di *Nature*. Dal confronto si vede che i valori che ho calcolato coincidono o quasi con quelli di *Nature*. Però c'è una notevole discrepanza per la sigma della media di Provenza, con una differenza di quasi dieci unità. Non so a che cosa attribuirlo. Non c'era alcun motivo di allargare la sigma perché per Provenza un chi-quadrato di 2,977 con 4 gradi di libertà produce un tranquillo valore di  $p = 0,56$ . Resta l'ipotesi che un calcolo sia stato sbagliato (da Zurigo o da me). Anche per Oxford abbiamo visto che per Provenza si calcola una sigma più piccola di quella riportata da *Nature*. Rifacendo i calcoli per l'intervallo di fiducia finale con i valori qui calcolati, si trova per Provenza una differenza non trascurabile, come vedremo al §22.



## §19. Calcolo delle medie entro laboratorio per Arizona

Il resoconto non dice e non lascia intendere niente sul metodo usato, ma mi sembra che si debba supporre che Arizona abbia usato il Metodo II con il moltiplicatore basato sul chi-quadrato. Non è possibile trovare un riscontro esatto perché sappiamo che le sigma dichiarate da Arizona nella Tabella 1 di *Nature* (quelle della nostra Tabella 1) non comprendono ancora la componente per l'incertezza nella misurazione del delta-C13, che è stata inglobata in un secondo tempo nella fase del calcolo delle medie. Non sappiamo quale sia l'incertezza attribuita da Arizona alla loro misurazione del rapporto C13/C12 e quindi non sappiamo quale sia il termine da aggiungere. Nemmeno sappiamo in quale modo, in concreto, il termine aggiuntivo sia stato incorporato. Possiamo solo presumere che l'introduzione dell'incertezza per il delta-C13 modifichi leggermente il risultato, ma non sappiamo di quanto. Questa carenza nei dati di cui disponiamo, ci impedisce di effettuare un controllo preciso confrontando da un lato i risultati per l'incertezza della media da noi calcolati con il Metodo II e dall'altro i risultati mostrati da *Nature*.

Non c'è bisogno che esplicitiamo il confronto fra le medie di *Nature* e quelle calcolate col Metodo II perché corrispondono perfettamente, al livello di arrotondamento usato da *Nature*. Cioè le medie sono le stesse medie ponderate che si calcolano col Metodo I, come ci si aspetta col Metodo II usando, per ciascun tessuto, un identico moltiplicatore per tutte le sigma dichiarate.

Nella Tabella 12 vediamo il confronto fra i valori di sigma della media forniti da *Nature* e quelli calcolati da noi col Metodo II.

Tabella 12 - Arizona, confronti per la sigma della media

	Metodo I	chi-q	gl	moltiplic.	Metodo II	Nature
Torino	17,04	8,671	3	1,7001	28,97	31
Nubia	19,70	9,055	4	1,5046	29,64	32
Tebe	19,89	22,342	4	2,3634	47,01	46
Provenza	20,42	16,730	4	2,0451	41,76	43

Nella prima colonna c'è la sigma della media ponderata quale calcolata con il Metodo I, cioè con la formula (4) e senza moltiplicatore. Nella quarta colonna c'è il moltiplicatore che si ottiene dividendo il chi-quadrato (seconda colonna) per i gradi di libertà (terza colonna) ed estraendo la radice quadrata. La quinta colonna, con i risultati calcolati col Metodo II, si ottiene moltiplicando la prima colonna per il moltiplicatore della quarta. L'ultima colonna ha i numeri quali pubblicati da *Nature*.

Nel fare il confronto, ripetiamo che va tenuto presente che il resoconto di *Nature* dice che Arizona non aveva incorporato nelle sigma dichiarate la componente per l'incertezza nella misurazione del delta-C13, una piccola correzione aggiuntiva che è stata inserita da Arizona al momento di fare le medie. Infatti si legge nel resoconto:

«The errors, which are quoted in Table 1 at the 1sd level (sd is standard deviation), include the statistical (counting) error, the scatter of results for standards and blanks, and the uncertainty in the  $\delta^{13}\text{C}$  determination (Arizona includes the  $\delta^{13}\text{C}$  error at a later stage, when combining subsample results.)»

Si vede che in tre casi, Torino, Nubia e Provenza, il risultato del nostro calcolo (quinta colonna) è di poco inferiore al valore della sigma della media fornito da *Nature* (sesta colonna). La piccola differenza potrebbe essere dovuta appunto al fatto che Arizona non aveva incluso l'incertezza nella misurazione del delta-C13 nelle sigma dichiarate e ne ha tenuto conto solo nel formare la media. Si può supporre che per l'incertezza del delta-C13 sia stato usato lo stesso valore per i vari tessuti.

Infatti il termine da aggiungere, in somma quadratica, ai valori della quinta colonna per avere la sesta colonna, sono circa uguali: 11,03 anni per Torino, 12,06 per Nubia, 10,25 per Provenza. Può essere che si trattasse sempre dello stesso identico termine perché le piccole differenze potrebbero essere dovute agli arrotondamenti nei numeri della sesta colonna. Infatti se si aggiunge alla quinta, in tutti i tre casi, sempre in quadratura, il termine 11, si trovano, dopo arrotondamento all'unità, gli stessi valori della sesta colonna.

Si può notare che i moltiplicatori sono diversi per i tre tessuti ma il termine da aggiungere per passare dalla quinta alla sesta colonna, cioè presumibilmente per tener conto dell'incertezza del delta-C13, è circa lo stesso per i tre tessuti. Questo fa supporre che in Arizona il termine per l'incertezza del delta-C13 sia stato aggiunto, dopo l'applicazione del moltiplicatore, sulle sigma già allargate. Se invece si aggiunge un termine per il delta-C13 inizialmente a ciascuna sigma dichiarata, poi si calcola il chi-quadrato e il moltiplicatore, i risultati per la sigma della media sono un po' inferiori.

Resta l'anomalia di Tebe, dove il valore calcolato non è un po' inferiore ma un po' superiore al valore di *Nature*. (Aggiungendo in quadratura al valore calcolato 47,01 un termine uguale a 11, si trova che invece di 46 ci dovrebbe essere 48,28.) Non se ne vede il motivo. Forse Arizona per Tebe ha usato un altro calcolo. (Si può notare, per quanto vale la cosa, che facendo il calcolo aggiungendo 11 in quadratura alle singole sigma dichiarate, prima di fare la media, si ottiene un moltiplicatore un po' più piccolo e risulta alla fine una sigma della media di 45,74, in accordo con il 46.) Oppure nel corso dei calcoli sono stati fatti arrotondamenti drastici (o si è partiti da valori non arrotondati un po' diversi da quelli da cui sono partito io prendendoli da *Nature*). Oppure c'è stato un errore di calcolo o di trascrizione, come abbiamo visto essere già successo per Zurigo/Provenza. Ma si vede quanto sia paradossale che dobbiamo fare le scommesse su quali siano stati i calcoli di *Nature*. I metodi di calcolo dovevano essere spiegati nel resoconto, non dovevano essere indovinati a tentoni da noi.

## §20. Le doppie pasticche di Arizona

Come abbiamo visto dalla Tabella 6, le sigma dichiarate di Arizona per Torino sono nettamente più basse rispetto agli altri tessuti per Arizona come rispetto agli altri laboratori.

Il motivo restò in origine sconosciuto finché, qualche anno dopo, Remi Van Haelst pubblicò informazioni che aveva ricevuto privatamente dal British Museum. A fornirglielo era stata Morven Leese, una delle autrici di *Nature*.

A quanto si può cercare di capire, sarebbe successo questo. Arizona aveva diviso il suo campione della Sindone (cioè la parte di campione utilizzata per la datazione) in quattro sottocampioni, Per ciascuno di questi, ottenne due pasticche di grafite, che inserì sullo stesso tamburo per la misurazione. Van Haelst non lo specifica, ma si può supporre che le due pasticche di ciascun sottocampione furono ottenute dopo processi di pulitura e di estrazione del carbonio eseguiti in comune e non separatamente. Quindi le due misurazioni non erano indipendenti perché le misurazioni furono fatte nella stessa seduta a confronto con gli stessi standard e gli stessi bianchi e, a quanto sembra, dopo l'identico processo di pulizia ed estrazione del carbonio.

Secondo quanto comunicato a Van Haelst dalla Leese, Arizona inviò a Londra otto risultati distinti per le quattro coppie di pasticche. Van Haelst ha pubblicato i dati e li ha ripetuti in diversi suoi articoli nel corso degli anni (per esempio in [14]). Non ci sono mai state smentite e quindi è da supporre che l'informazione sia corretta. Nella Tabella 13 riportiamo nelle prime due colonne i risultati per le quattro coppie.

Tabella 13 – Arizona/Torino, otto risultati

606 ± 41	574 ± 45	591 ± 30
753 ± 51	632 ± 49	690 ± 35
540 ± 57	676 ± 59	606 ± 41
701 ± 47	701 ± 49	701 ± 33

Nella terza colonna c'è la media ponderata (col Metodo I) dei due risultati di ciascuna coppia. Questi della terza colonna sono i valori pubblicati su *Nature* per Arizona/Torino.

A quanto sembra, sempre secondo le indiscrezioni pubblicate da Van Haelst, a Londra si ritenne che i dati di ogni coppia non fossero indipendenti e andassero fusi. Quindi i numeri indicati in Tabella 1 di *Nature* per Arizona/Torino non si riferiscono a singole datazioni ma alle medie fra due datazioni per le due pasticche di ciascun sottocampione. Come si vede, i valori di sigma nella terza colonna diventano inferiori rispetto alle prime due colonne. Naturalmente, calcolando la media di due datazioni si ottiene una sigma della media che è ridotta, rispetto alle sigma delle singole datazioni, per un fattore di circa radice di 2 (1,41).

Il resoconto di *Nature* non accenna affatto a questa circostanza delle doppie pasticche. È una mancanza grave perché chi legge crede che i valori indicati si riferiscano a singole datazioni mentre invece si tratta di medie di due datazioni.

La mancata comunicazione è deplorabile, ma le cose non cambiano in pratica. La media ponderata, calcolata dagli otto valori o dai quattro valori, resta praticamente identica. Calcolando sugli otto dati, si ottiene 645,93; calcolando sui quattro dati di *Nature* si ottiene 646,44. *Nature* arrotonda a 646, che corrisponde a entrambi.

Se Arizona avesse deciso, come probabilmente sarebbe stato opportuno, di non fornire dati separati per ogni coppia di pasticche ma di congregarli all'origine, si sarebbe fatta la media di tutte le misurazioni sulle due pasticche, e sarebbe risultata in pratica la stessa media. Per la sigma dichiarata, si sarebbe dovuto fornire un valore più alto. Per ironia, con valori più grandi della sigma dichiarata sarebbe stato probabilmente superato il test del chi quadrato. La cosa quindi non ha influenza sulle medie ma ha influenza sulle sigma dichiarate. Dato che per Arizona è stata calcolata la media entro laboratorio con il moltiplicatore, ci sarebbe stato un moltiplicatore più piccolo (un chi-quadrato più piccolo) con sigma dichiarate più grandi. Non so se alla fine si sarebbe trovata una sigma della media un po' più grande o un po' più piccola, ma la differenza non dovrebbe essere rilevante.

### §21. Calcolo della media dei tre laboratori per Nubia, Tebe e Provenza

Veniamo alla seconda fase dei calcoli, quella per ottenere, per ciascun tessuto, il risultato finale espresso da un unico intervallo di fiducia, conglobando i tre risultati per le medie di ciascun laboratorio. Per ciascun tessuto, *Nature* ha applicato il test del chi-quadrato alle tre medie dei tre laboratori.

Per Nubia, Tebe e Provenza il test chi-quadrato è ampiamente superato. La media delle medie, con la relativa sigma, è stata calcolata con la media ponderata (Metodo I). Si trova così il risultato finale per l'intervallo degli anni BP (intervallo al 68% di fiducia se si indica più/meno una sigma) I risultati sono quelli forniti da *Nature* nella sua Tabella II (sotto “Weighted mean”) e ripetuti nella sua Tabella 3 dove sono mostrati anche i corrispondenti intervalli di data calibrata, questi sia al 68% sia al 95% di fiducia.

## §22. Calcolo della media dei tre laboratori per Torino

Per Torino il test del chi-quadrato, per le differenze fra le medie dei tre laboratori, non è superato a causa delle età più antiche di Oxford. Le autrici ricorrono allora al Metodo III.

Il calcolo si esegue nel seguente modo. Si fa la media semplice (non ponderata) delle tre medie di Arizona, Oxford e Zurigo. Per la sigma della media delle medie, si procede così. Si calcola la dispersione delle tre medie e si divide per la radice di 3, secondo la formula (7). Si trova 31

Per calcolare l'intervallo di fiducia col Metodo III, come abbiamo detto, si usa la *t* di Student. Non so se sia valida o più o meno arbitraria la scelta di *Nature* di usare la *t* con 5 gradi di libertà, tenendo conto anche del numero totale di datazioni. La *t* per 5 gradi di libertà al 95% è 2,57, che *Nature* arrotonda a 2,6. (Per 2 gradi di libertà sarebbe 4,30, cioè alquanto più grande). Quindi per l'intervallo di fiducia in anni BP al 95% la semiampiezza è il prodotto di 31 per 2,6, cioè 80.

Si può essere perplessi per l'uso del Metodo III con soli tre valori sui quali calcolare la dispersione, ma nel nostro caso c'è la giustificazione che con questo metodo si trova l'intervallo di fiducia più ampio rispetto ad altri metodi, cioè in pratica si ha la risposta più prudente.

Nella Tabella 14 sono messe a confronto le ampiezze per l'intervallo di fiducia in anni BP al 95% calcolate con diversi metodi.

Tabella 14 - Ampiezza dell'intervallo di fiducia

Torino	(64)	78	160
Nubia	64	67	23
Tebe	80	103	84
Provenza	80	79	104

Nella prima colonna c'è il metodo usato per Nubia, Tebe e Provenza (Metodo I dove per le medie di Arizona si usano le sigma allargate col Metodo II). Per questi tre tessuti i valori riportati per l'ampiezza dell'intervallo sono appunto quelli ai quali perviene *Nature*. Il valore per Torino, in parentesi, non sarebbe valido.

Nella seconda colonna ho usato il metodo III applicato direttamente ai valori cumulati di tutti i tre laboratori. Ho usato ( $n - 1$ ) come numero di gradi di libertà (usando  $n - 3$  gli intervalli sarebbero leggermente più ampi, per esempio per Torino 80 invece di 78).

Nella terza colonna c'è il metodo usato per Torino (Metodo I, con Metodo II per Arizona, per la prima fase, poi Metodo III per la seconda fase), e l'ampiezza dell'intervallo per Torino è appunto quella di *Nature*. Si vede che per Nubia si avrebbe un intervallo molto stretto.

Ho provato a usare altri due metodi per Torino. Usando il Metodo II anziché il Metodo III nella seconda fase, si trova un intervallo di 115 anni. Usando il Metodo I sui valori cumulati dei tre laboratori (dopo avere allargato le sigma di Arizona col moltiplicatore, ciò che è ampiamente sufficiente per far superare il test del chi-quadrato per l'insieme dei dodici valori), si trova un intervallo di 63 anni.

Aggiungo che per Provenza, usando per le incertezze delle medie i valori quali da me calcolati nel §18, si otterrebbe un intervallo più stretto, cioè 66 invece di 80.

Da questo balletto di numeri si vede che, come dicevamo, i risultati di questi metodi vanno considerati con molta elasticità. Il metodo usato da *Nature* per Torino è quello che produce l'intervallo più ampio e questo è un buon motivo per preferirlo.

Infine guardiamo al passaggio alle date di calendario per Torino. Dopo la calibrazione, *Nature* fornisce un intervallo al 95% che è diviso in due segmenti: dal 1262 al 1312 e dal 1353 al 1384 d.C. Il secondo segmento, che avrebbe comunque una probabilità molto bassa, nasce solo dal fatto che l'andamento della curva di calibrazione per il secolo XIV è irregolare. Come si vede dal grafico che è riprodotto su *Nature*, la curva, invece di scendere regolarmente, attorno alla metà del secolo risale

per un tratto. Ne consegue che un reperto della fine del Trecento, misurato oggi, ha la stessa proporzione di C14 come un reperto dell'inizio del secolo, e quindi è indistinguibile come età radiocarbonica. Naturalmente, se la Sindone oggi a Torino è la stessa che fu esposta a Lirey attorno alla metà di quel secolo, il secondo segmento dell'intervallo non è da considerare. Il primo intervallo, di soli 50 anni, è così stretto perché in quel tratto la curva di calibrazione è piuttosto ripida.

In conclusione, abbiamo visto che il resoconto di *Nature*, per quanto riguarda la statistica, soffre di qualche difetto di comunicazione, per così dire. In particolare, non è stata evidenziata la circostanza delle doppie pastiche per Arizona/Torino e non è stato esplicitato il metodo di calcolo delle incertezze delle medie per i quattro tessuti di Arizona. Per il resto, i metodi impiegati sono sostanzialmente corretti.

## II – Le accuse dei sindonologi

In questa seconda parte esaminiamo le critiche che sono state avanzate dai sindonologi sulla trattazione statistica del resoconto di *Nature*.

Il primo sindonologo a criticare la statistica della datazione è stato probabilmente il belga Remi Van Haelst (per esempio [14]) ma qui ci limitiamo alla scena italiana, dove il pioniere è stato Ernesto Brunati dagli anni '90. Il suo punto di vista si può trovare esaurientemente esposto in un articolo del 2005 [2]. Brunati collocava le critiche di tipo matematico nell'ambito di uno scenario fantasioso in cui accusava di frode i responsabili della datazione, da Michael Tite ai direttori dei tre laboratori. Essi avrebbero effettuato una sostituzione del campione della Sindone e avrebbero datato al suo posto un altro tessuto. Le accuse di complottismo avanzate da Brunati sono così inverosimili che non vale la pena di considerarle e qui ci limiteremo agli aspetti statistici.

Le tesi di Brunati sono rimaste per qualche anno in sordina. Poi la linea di attacco basata sulla statistica è stata rilanciata da Marco Tosatti con un libro del 2009 [11]. Tosatti si è avvalso della collaborazione di due professori di statistica della Sapienza di Roma, Livia De Giovanni e Pierluigi Conti. I due docenti hanno implicitamente dato sostegno a Tosatti che può avvalersi della loro presunta autorevolezza nel propagandare le sue tesi.

In generale gli altri sindonologi sono rimasti a guardare e non hanno rilanciato in prima persona le accuse, forse perché non le hanno capite, ma ci sono anche stati quelli, come Giulio Fanti ed Emanuela Marinelli, che le hanno ripetute nei loro libri recenti. Comunque le tesi hanno avuto risonanza sui giornali o su internet in seguito ad articoli o interviste pubblicati da Tosatti, Fanti e Andrea Tornielli.

Quando nel seguito ci riferiamo ai “sindonologi”, intendiamo questi pochi autori che si sono espressi direttamente. Va comunque notato che non risulta che gli altri sindonologi si siano dissociati da queste accuse basate sulla statistica. Quindi tali accuse sono entrate a far parte del repertorio dei sindonologi e si possono considerare approvate dalla categoria nel suo insieme.

Esaminiamo le posizioni dei sindonologi elencando una serie di punti.

### §23. *I sindonologi non sembrano capire che la sostanza del risultato della datazione non dipende dalle analisi statistiche eseguite.*

I sindonologi si sono sforzati di attaccare il risultato medievale della datazione cercando difetti nel trattamento statistico pubblicato su *Nature*, cioè nei calcoli eseguiti per combinare i dodici risultati e ottenere un unico intervallo di fiducia finale. Ma il loro tentativo è destinato in partenza a fallire perché la sostanza del risultato non dipende dai calcoli statistici che qualcuno ha fatto. Il risultato

consiste in dodici datazioni effettuate in tre laboratori, che tutte sono racchiuse nell'arco di duecento anni radiocarbonici convenzionali, un intervallo soddisfacente per quanto ci si poteva aspettare all'epoca. Traducendo in date di calendario calibrate, si trova il periodo immediatamente anteriore alla prima comparsa della Sindone nella storia. I laboratori hanno contemporaneamente datato altri tre tessuti trovando i risultati previsti. Il discorso è finito qui.

Per un desiderio di completezza formale che era comunque superfluo, gli autori di *Nature* hanno voluto combinare assieme i dodici risultati per fornire un elegante unico intervallo di fiducia complessivo. I sindonologi si sono dedicati a cercare errori nei calcoli. Non ci sono errori significativi ma, anche se ce ne fossero, non potrebbero cambiare la sostanza del risultato. Con dodici datazioni attorno al 1300 d.C., nessuna analisi matematica potrebbe realizzare il sogno dei sindonologi, che è quello di datare la Sindone al primo secolo.

Alcuni sindonologi pretendono che con la statistica si possa dimostrare che il materiale dei vari frammenti datati non fosse omogeneo nel contenuto di C14, per esempio a causa di inquinamenti o rammendi, ma abbiamo visto che la variabilità fra i dodici campioni è nei limiti di quanto ci si poteva aspettare dagli errori sperimentali e non implica che ci siano reali differenze fra i campioni.

Con i loro sforzi, i sindonologi hanno solo dimostrato di non avere compreso i metodi di calcolo impiegati dagli autori del resoconto di *Nature*.

#### §24. *I sindonologi credono che gli autori di Nature per combinare i risultati abbiano usato un solo metodo.*

Invece sono stati usati tre metodi, come abbiamo mostrato nella prima parte dell'articolo. I sindonologi non se ne sono accorti.

I sindonologi parlano a volte di “Metodo di Ward e Wilson” per indicare l'unico metodo che conoscono. Intendono quello che nella prima parte abbiamo indicato come Metodo I. Si tratta semplicemente del calcolo della media ponderata dove i pesi sono gli inversi delle varianze. È improprio chiamarlo metodo di Ward e Wilson perché il metodo non è stato inventato da loro. È stato solo da loro consigliato in un articolo del 1978. [15]

I due autori australiani, l'archeologo Graeme K. Ward e l'esperta di statistica Susan R. Wilson, all'inizio dell'articolo (p. 20-21) presentano il metodo della media ponderata e aggiungono il consiglio di fare la verifica dell'omogeneità delle varie datazioni col test del chi-quadrato. Gli autori di *Nature* citano quell'articolo solo in riferimento al test del chi-quadrato.

Quello suggerito da Ward e Wilson è un metodo standard per combinare diverse distribuzioni di probabilità che hanno varianze diverse. Richiede che siano verificate condizioni stringenti: che le distribuzioni siano normali, che abbiano tutte lo stesso valore centrale, che abbiano varianze diverse fra loro ma tutte note. È illusorio pensare che, specialmente col livello di qualità prevalente negli anni '80, tutte queste condizioni siano verificate per diverse misurazioni radiocarboniche dell'età di un reperto. Comunque in teoria il metodo sarebbe valido ed è ben conosciuto, come possiamo vedere per esempio da una pagina di Wikipedia dedicata alla media ponderata. [16] Nel paragrafo “Dealing with variance” si trovano le formule che abbiamo mostrato per il Metodo I (§14). (Lì si parla di varianze e ricordo che la varianza è il quadrato della deviazione standard, cioè della sigma.) Ward e Wilson nel loro articolo usano il metodo applicandolo direttamente a tutti i dati cumulati ottenuti da diversi laboratori, mentre *Nature* ha operato in due fasi, prima calcolando le medie entro ciascun laboratorio, poi facendo la media delle medie.

#### §25. *I sindonologi credono che le medie entro laboratorio di Arizona siano state calcolate con il Metodo I e usando le sigma dichiarate.*

Invece le medie di Arizona, per tutti e quattro i tessuti, sono state calcolate allargando le sigma dichiarate. Si può supporre che sia stato usato un moltiplicatore come abbiamo illustrato per il

Metodo II (§16). Ai sindonologi va riconosciuto che il resoconto di *Nature* non dice niente sul metodo di calcolo delle medie di Arizona. In realtà non ho garanzia che il metodo usato sia proprio quello che ho indicato come Metodo II e quindi è opportuno spiegare i motivi che inducono a ipotizzarlo.

Dai numeri pubblicati su *Nature*, si vede che le medie entro laboratorio di Arizona, per ciascuno dei quattro tessuti, sono le stesse che si ottengono con le medie ponderate con pesi uguali agli inversi delle varianze dichiarate, cioè con il Metodo I. Invece le relative incertezze (le sigma delle medie) sono accresciute rispetto a quelle che si otterrebbero con le varianze dichiarate. Questo fornisce una indicazione sul metodo usato. È stato applicato un moltiplicatore alle sigma dichiarate, lo stesso per tutte le sigma (per ciascun tessuto). In questo modo la media ponderata resta la stessa perché i rapporti fra le varianze restano uguali, mentre la sigma della media viene aumentata perché moltiplicata per il moltiplicatore. Si noti infatti che, con il Metodo II, nella formula (3) del Metodo I (§14) per il calcolo della media tutte le sigma vengono moltiplicate per lo stesso fattore che così compare sia a numeratore che a denominatore e si elide. Nella formula (4) per il calcolo della varianza, o della sigma, il fattore compare solo al numeratore e non si elide.

Se proviamo a fare altre ipotesi, diverse da quella del moltiplicatore, troviamo che la media risulterebbe (leggermente) diversa da quella calcolata col Metodo I. Si potrebbero immaginare altri criteri per allargare le sigma dichiarate, per esempio aggiungendo uno stesso termine in quadratura a ciascuna sigma, ma allora i rapporti fra le varianze non sarebbero più gli stessi. (Però se le sigma dichiarate sono poco diverse fra loro, la differenza rispetto alla media del Metodo I potrebbe essere così piccola da essere inglobata nell'arrotondamento.) Oppure si potrebbero ignorare del tutto le sigma dichiarate usando il Metodo III, ma allora si farebbe la media semplice e non ponderata.

Invece abbiamo visto che per i quattro tessuti di Arizona la media presentata è la stessa che si ottiene col Metodo I. Va detto che le medie sono presentate nel resoconto arrotondate all'unità e una piccola differenza nei decimali non sarebbe rilevabile, ma il fatto che per tutti e quattro i tessuti ci sia coincidenza fa pensare che il metodo usato sia quello del moltiplicatore. Comunque le medie semplici sono nei quattro casi diverse rispetto alle medie presentate, anche lasciando un margine per gli arrotondamenti. Assumiamo quindi che sia stato usato un moltiplicatore delle sigma dichiarate, sempre lo stesso per tutti i risultati di ciascun tessuto.

Quale moltiplicatore può essere stato usato? La scelta ovvia è di tener conto del valore del chi-quadrato, che è appunto quello che denuncia che le sigma dichiarate sono sottostimate. Più è grande il chi-quadrato, più bisogna aumentare il moltiplicatore, Ma il chi-quadrato aumenta di suo anche all'aumentare dei gradi di libertà. Quindi una scelta opportuna è di sostituire le varianze dichiarate con le varianze ottenute moltiplicando per il chi-quadrato e dividendo per il numero dei gradi di libertà:

$$\sigma^2 \rightarrow \sigma^2 \frac{\chi^2}{n-1}$$

Guardando una tavola del chi-quadrato, e scorrendo in orizzontale lungo una riga per un dato numero di gradi di libertà, si vede che questo rapporto si porta vicino all'unità quando il valore del chi-quadrato è centrale (con il livello di significatività  $p$  attorno a 0,50), cioè quando non c'è bisogno di allargare le sigma. Se il chi-quadrato aumenta, portandosi verso la coda della distribuzione dove  $p$  scende verso lo zero, allora le sigma dichiarate sono sottostimate e occorre allargarle, e infatti il rapporto fra il chi-quadrato e il numero di gradi di libertà supera uno e sale tanto più quanto più  $p$  scende. Se il chi-quadrato invece è piccolo e si porta verso l'altra coda della distribuzione (con  $p$  che sale verso uno), allora le sigma sono sovrastimate e il rapporto diventa inferiore a uno, quindi applicando il moltiplicatore le sigma si riducono.

Appunto questo moltiplicatore viene indicato nella già vista pagina di Wikipedia sulla media ponderata. La sezione successiva a quella già citata ha il titolo "Correcting for over- or under-dispersion". Vi si legge:

«Weighted means are typically used to find the weighted mean of experimental data, rather than theoretically generated data. In this case, there will be some error in the variance of each data point. Typically experimental errors may be underestimated due to the experimenter not taking into account all sources of error in calculating the variance of each data point. In this event, the variance in the weighted mean must be corrected to account for the fact that  $\chi^2$  is too large. The correction that must be made is...» [16]

E segue la formula che abbiamo indicato per il moltiplicatore e che abbiamo usato nella prima parte col Metodo II. Si vede che la situazione descritta da Wikipedia è proprio come quella per i dati di Arizona.

Aggiungo che ci potrebbero essere varianti nel metodo. Nel denominatore del moltiplicatore si potrebbe usare  $n$  anziché  $(n-1)$ . Nel calcolo del chi-quadrato si potrebbe usare la media semplice anziché ponderata. Comunque le differenze sarebbero piccole. C'è poi da considerare che Arizona aveva coppie di risultati ottenuti con pasticche gemelle prodotte con lo stesso sottocampione. Per Torino c'erano quattro coppie. Per gli altri tre tessuti, a quanto desumibile dal resoconto, c'era una coppia per ciascuno. Non sappiamo esattamente se nel calcolo delle medie si sia partiti accorpando i due risultati di una coppia, come sembrerebbe sia stato fatto per Torino, oppure tenendo i due risultati distinti come sembrerebbe per gli altri tre tessuti. Insomma è possibile che il metodo di calcolo applicato da Arizona non sia esattamente quello da noi indicato come Metodo II, ma le eventuali differenze dovrebbero essere trascurabili.

La situazione è complicata dal fatto che le sigma dichiarate di Arizona non tengono ancora conto dell'incertezza nella misurazione del delta-C13 e non conosciamo l'entità della correzione.

È ovviamente da rilevare che il laboratorio di Arizona avrebbe dovuto spiegare il metodo usato per calcolare le medie dei propri risultati, per la Sindone come per gli altri tessuti. Non avendolo fatto nel resoconto di *Nature*, poteva comunicarlo negli anni successivi quando i sindonologi attirarono l'attenzione sul presunto errore commesso nel calcolare le medie. Dopo più di venti anni dalla datazione, non mi risulta che Arizona (né Londra) abbia mai fornito una spiegazione. Io ho cercato di indovinare quale metodo abbiano usato e mi sembra che la mia ipotesi sia giustificata, ma non posso garantire che sia così e dovremo continuare ad aspettare che Arizona dica una parola definitiva. (Ma si può temere che oggi a Tucson non ci sia più nessuno che abbia memoria o documentazione di quale fosse il metodo adottato.)

## §26. *I sindonologi affermano con sicurezza che gli autori di Nature hanno commesso un errore nel calcolo della sigma della media per Arizona/Torino: 31 al posto di 17.*

Infatti loro calcolano con il Metodo I e trovano 17, mentre il valore di *Nature* è 31. Non pensano che se *Nature* fornisce un numero diverso, può essere perché ha eseguito i calcoli con un altro metodo. Questo non dovrebbe sembrare strano ai sindonologi perché loro stessi fanno notare che il test del chi-quadrato non viene superato per Arizona e allora dovrebbero sapere che ciò impedisce di accettare il valore della sigma della media calcolato col Metodo I. Dovrebbe anche essere intuitivo che, se le sigma dichiarate erano troppo strette, il valore 17 sarebbe stato sbagliato, mentre un valore allargato come 31 poteva andar bene.

Rimando alla prima parte (§19) per il calcolo eseguito con il moltiplicatore delle sigma. Ricordo che il calcolo fornisce un valore di 29 e che ho attribuito la piccola differenza, rispetto a 31, al fatto che, come specificato nel resoconto, Arizona nelle sigma dichiarate non aveva ancora tenuto conto dell'incertezza nella misurazione del delta-C13.

Il primo in Italia a “scoprire” l'errore del 31 al posto del 17 è stato Ernesto Brunati, nel 1997 se non prima. Un quadro completo delle sue accuse si trova nel citato articolo del 2005. [2]

La presenza dell'errore è stata poi propagandata da Marco Tosatti che la denuncia nel suo libro [11] dove riporta ampi passi dell'articolo di Brunati. Tosatti ne ha riparlato anche in articoli o interviste su giornali o internet.



Il secondo capitolo del libro di Tosatti si intitola “Un errore minuscolo ma colossale” (p. 14). Vediamo qualche frase.

«Ma c'è di più. Perché anche nell'unico elenco di cifre che i laboratori – e il British Museum – hanno reso di dominio pubblico, e cioè l'articolo pubblicato su “Nature”, emerge un errore grossolano, tale da inficiare la validità dell'esame. Un errore di calcolo (!) compiuto da un laboratorio, quello di Tucson, in Arizona, di cui non si sono accorti né il British Museum, né i responsabili dell'Istituto Metrologico Colonnetti di Torino. Un errore che – nelle parole di uno degli esperti di statistica che abbiamo consultato – non dovrebbe/potrebbe fare nemmeno uno studente del secondo anno di matematica.» [11, p. 15]

Questo esperto di statistica offende la sua categoria se pensa che qualche suo collega abbia potuto commettere un errore simile.

Ancora Tosatti più avanti nel libro:

«Stupisce, e non può non stupire, che un rapporto [di *Nature*] firmato da una ventina di professori, contenga un errore di calcolo dalle conseguenze così gravi; e che questa irregolarità non sia stata scoperta, messa in luce e contestata molto tempo fa da chi dovrebbe essere più interessato alla Sindone, e cioè dalla diocesi di Torino che, invece, all'epoca del cardinale Ballestrero, è parsa molto remissiva nella conduzione di tutta questa vicenda. È certamente una situazione che autorizza grandi perplessità.» [11, p. 29]

Sempre Tosatti in una intervista all'agenzia cattolica internazionale Zenit [9]:

«Ma anche solo esaminando i dati pubblicati su *Nature*, un ingegnere di Milano, Ernesto Brunati, si è accorto che c'era qualche cosa che non andava. Ho chiesto di rifare i calcoli a due professori di matematica e statistica della Sapienza, che non c'entrano nulla con il mondo della Sindone. Livia De Giovanni e Pierluigi Conti, che hanno confermato: c'era un errore di calcolo, tale da inficiare la validità dell'esame.»

Quindi coinvolge direttamente i due professori.

Ancora Tosatti in un articolo su *La Stampa* all'epoca dell'ostensione del 2010 [12]:

«Ma qualche tempo dopo, un appassionato della Sindone, l'ingegner Ernesto Brunati, si prese la briga di rifare i conti; e scoprì un errore di calcolo, inspiegabile ma fondamentale.»

Vediamo ora una frase di Giulio Fanti, intervistato da Andrea Tornielli sul *Giornale* [10]:

«Com'è potuto accadere un errore di questo genere?»

«“Diversi studiosi hanno dimostrato già da tempo che i risultati della datazione, in base al test statistico cosiddetto di Pearson [del chi-quadrato], hanno una probabilità superiore al 95% di non corrispondere a quelli della Sindone. È stato inserito nella formula un numero sbagliato, che ha falsato il risultato finale, un 31 è stato sostituito con un 17. Questo farebbe pensare persino a una manomissione finalizzata a ottenere il risultato desiderato”».

Ed ecco una invettiva di Brunati nell'articolo del 2005 [2]:

«Ma c'è un altro aspetto della vicenda che non credo debba essere trascurato. La datazione è stata una prova svoltasi di fronte a tutto il mondo e chi l'ha effettuata non ha lavorato privatamente, ma a nome di illustri Università e, specialmente, a nome di una istituzione famosa come il British Museum. È mai possibile che simili autorità accademiche accettino che loro dipendenti abbiano a asserito, a loro nome, che il risultato di quella famosa elementare espressione aritmetica fosse

uguale a 31 anziché a 17? E che, per giunta, lo abbiano anche pubblicato, sempre a loro nome, su una rivista considerata famosa e scientifica come *Nature*?

«E, visto che quell'errore comporta l'annullamento di una prova che tutti considerano importante, è mai possibile che quelle autorità accademiche non intervengano?»

I sindonologi non si accorgono che ogni volta che denunciano il presunto errore, dimostrano di essere loro stessi a sbagliarsi.

### §27. *Qualche sindonologo lancia l'accusa che il 31 al posto del 17 sia frutto di una alterazione deliberata e fraudolenta.*

Brunati non si è limitato a denunciare l'errore da lui "scoperto", ma ha avanzato l'accusa che l'alterazione del numero, da 17 a 31, sia stata intenzionale e fraudolenta. Lo scopo della falsificazione? Far credere che il test del chi-quadrato, nel successivo confronto fra i tre laboratori, venisse superato, cioè arrivasse a produrre per il livello di significatività  $p$  un valore di almeno 0,05. Naturalmente con un valore maggiore per la sigma della media di uno dei laboratori (Arizona), diminuisce il valore del chi-quadrato per il confronto fra i tre laboratori. (Ricordo che il valore di  $p$  aumenta al diminuire del valore del chi-quadrato.)

Quanto sia maldestra questa accusa, lo si capisce considerando che gli autori di *Nature* non hanno considerato superato il test del chi-quadrato per il confronto fra i tre laboratori, tanto che per il calcolo della media delle medie dei tre laboratori non hanno proseguito con il Metodo I ma hanno usato il metodo III (§22). Del resto, anche usando il 31 non viene superato il test del chi-quadrato e la  $p$  si ferma a 4,17%.

Ma Brunati scrive [2]:

«La verifica col *chi quadro* richiedeva, dapprima, il calcolo del valore dello stesso *chi quadro* dal quale si passava, poi, al cosiddetto *significance level*. La prova si intendeva superata se quest'ultimo, una percentuale, era almeno pari al 5%. I risultati riportati sulla tabella del rapporto erano stati ottenuti, ovviamente, partendo dal responso sbagliato di Tucson Arizona,  $646 \pm 31$ . Essi davano un *chi quadro* di 6,4 ed un *significance level* pari al 5%, il minimo richiesto. Un valore tale, comunque da far considerare la prova superata.»

(Brunati usa sempre l'espressione inglese *significance level* invece dell'italiano "livello di significatività", quello che noi indichiamo con  $p$ .)

In effetti su *Nature* il livello di significatività è arrotondato a 5% e Brunati cade nell'equivoco di credere che il test fosse dagli autori considerato superato. Anzi, l'ingresso di Brunati nelle dispute sulla statistica fu nel 1993 quando accusò gli autori di *Nature* di avere imbrogliato perché indicavano un livello di significatività del 5% mentre era un po' inferiore. Ma *Nature* non voleva imbrogliare perché indica anche il valore del chi-quadrato, anzi lo arrotonda per eccesso. Comunque in questo tipo di situazione, anche un valore di  $p$  attorno al 5% o di poco superiore induce a prudenza.

Brunati continua, di seguito al brano già citato:

«Quando, però, ho rifatto il calcolo partendo dalla media corretta,  $646 \pm 17$ , il *chi quadro*, diventava 9,8 mentre il *significance level* si riduceva ad un modestissimo 1,04%, decisamente troppo basso perché la prova potesse essere considerata superata.

«Questo mi ha fatto finalmente capire quale fosse il vero motivo che li aveva spinti a cambiare quel risultato. Hanno dovuto ricorrere a quel falso per superare, almeno sulla carta, quella prova che avrebbe consentito loro di far quella media fra i 646 anni di Tucson, i 750 anni di Oxford ed i 676 anni di Zurigo, la media che avrebbe concluso la datazione. Chiaro, quindi, che non si trattasse di un errore.»

Brunati ribadisce in un articolo del 2006 [3]:

«Interessante, a tale proposito, notare che la tab. 2 di *Nature* riporta tutti i *significance levels* raggiunti facendo quella datazione. Da essi risulta che, mentre per i campioni di controllo si erano raggiunti valori pari al 90, 50 e 30%, per la datazione della Sindone si era proprio al minimo; si raggiungeva, cioè, un *significance levels* pari ad appena il 5%. Grazie al quale, però, il British Museum poteva ritenere ammissibile far la media e concludere la datazione.

«Quando, però, ho rifatto il calcolo, considerando, invece del  $646 \pm 31$ , sbagliato, del quale aveva tenuto conto il British Museum, il  $646 \pm 17$  corretto, mi sono accorto che quel *significance level* si riduceva dal 5 a circa l'1 %. In altre parole, facendo i conti giusti, risultava che la prova col chi quadro non poteva venir superata, per cui, non essendo lecito far la media finale, non sussistevano nemmeno le condizioni per concludere la datazione.

«Risultava pertanto evidente il motivo per cui avevano cambiato quel 17 in 31. Dovevano far apparire superata la prova col chi quadro, che avrebbe consentito l'esecuzione della media finale e, quindi, di portar a termine la datazione. Un ulteriore motivo per far considerare quella sostituzione non un errore accidentale ma un provvedimento preso di proposito e per capire l'importanza che aveva avuto sull'esito della datazione.

«A proposito di *significance levels*. Quando ho scoperto che quel  $\pm 17$  era diventato  $\pm 31$ , mi sono anche chiesto: come mai proprio  $\pm 31$ ? Poi, mi sono accorto che quella era la tolleranza minima alla quale si doveva arrivare per avere un *significance level* almeno uguale al 5%, un *significance levels* tale da consentir di dire che la prova del chi quadro era stata superata. Anche se, rifacendo in modo più attento quei conteggi, se si teneva conto di un altro risultato che *Nature* forniva, cioè il valore del chi quadro pari a 6,4, ci si sarebbe accorti che nemmeno quel 5% era corretto. In realtà, era solo un 4%. Ma, visto il resto, questo, lo possiamo considerare un dettaglio irrilevante.»

§28. *Qualche sindonologo cade nell'equivoco di credere che, se la media di Arizona è quella che si ottiene col Metodo I, anche la sigma della media dovrebbe essere la stessa ottenuta con quel metodo.*

Brunati crede di trovare un altro motivo per pensare che la sostituzione di 31 al posto di 17 sia stata intenzionale e fraudolenta. (Nel leggere questo e successivi passi di Brunati, si tenga presente che lui usa il termine “tolleranza” per indicare l'incertezza o sigma. Inoltre Brunati scrive sempre, sbagliando, “media delle tolleranze” quando dovrebbe scrivere “tolleranza della media”.) Dall'articolo del 2005 [2]:

«Parlo di sicura alterazione del risultato escludendo un possibile errore accidentale anche per il motivo seguente. Se considerate la formula che porta alla media delle età (quella che, nel caso particolare, ha portato al 646) noterete che essa praticamente consiste in una frazione con l'espressione

$$\left( \frac{1}{30^2} + \frac{1}{35^2} + \frac{1}{41^2} + \frac{1}{33^2} \right)$$

che svolge la funzione di denominatore. Essa dà, come risultato 0,003440595 ed è la stessa che porta alla determinazione della media delle tolleranze. Ne viene che, se il British Museum ha ottenuto il valore corretto della media delle età (646) deve, per forza, avere ottenuto, quale risultato del calcolo di quell'espressione, il numero 0,003440595, la radice quadrata dell'inverso del quale non è altro che quel 17 che rappresenta il valore corretto della media delle tolleranze. Dando come media delle età quel 646, quindi, implicitamente, si ammette che la media delle tolleranze sia 17 e non 31. Quel risultato che troviamo in tab. 2 del rapporto ufficiale, quel  $646 \pm 31$ , quindi, non può

sussistere. E non può essere un errore di calcolo, ma è un risultato volutamente alterato, è un vero e proprio dolo.»

Se riguardate le formule (3) e (4) (§14), vedete che quella sommatoria degli inversi delle varianze, esplicitata da Brunati, compare sia nella formula per la media ponderata, sia nella formula per la sigma della media. Come abbiamo detto al §16, col Metodo II, usando lo stesso moltiplicatore per tutte le sigma dichiarate, nella formula per la media il moltiplicatore compare sia al numeratore che al denominatore e si elide, mentre nella formula per la sigma della media il moltiplicatore non si elide. Quindi il risultato per la media è lo stesso sia col Metodo I (senza moltiplicatore) sia col Metodo II (con moltiplicatore). Invece la sigma della media col Metodo II risulta aumentata, e precisamente moltiplicata per il moltiplicatore stesso. Appunto questa circostanza, che la media rimane la stessa mentre la sigma della media viene aumentata, ci è servita come indizio per supporre che fosse stato usato il Metodo II per la media di Arizona. Brunati non lo sa e, per il fatto che il valore della media è lo stesso da lui trovato col Metodo I, crede che anche la sigma della media dovrebbe essere la stessa. Ne trae motivo per confermarsi nella sua tesi della alterazione deliberata del numero.

Anche Fanti cade nello stesso equivoco. Nel suo libro scrive [8, p. 159-160]:

«Lo sbaglio forse più evidente nell'analisi statistica del rapporto di *Nature* è forse quello relativo all'assegnazione dell'errore [sigma della media], in riferimento ai dati del laboratorio di Tucson. I quattro valori delle date risultanti [nota 49] furono elaborati per ottenere il valore di  $646 \pm 31$  anni anziché  $646 \pm 17$  anni, ma il valore 646 si può ottenere solo se si inserisce nella formula il valore 17 e non 31! Questo fa pensare che forse qualcuno abbia cercato di manomettere i risultati per arrivare al dato desiderato.»

Nella nota 49 a fondo pagina, Fanti esplicita le formule per la media e per la sigma della media. Dice:

«Chi scrisse il rapporto non può quindi avere calcolato erroneamente 31 anziché 17 perché la media pesata delle età fu calcolata con la relazione seguente dove ricompare il valore derivato dai reciproci degli errori misurati che corrisponde al valore 17.»

*§29 I sindonologi sorvolano sul fatto che un presunto errore nella sigma della media si riscontra anche per gli altri tessuti di Arizona. Solo Brunati lo dice, ma ne deduce un ulteriore motivo per la sua accusa di alterazione deliberata.*

Come sappiamo (§19), Arizona fornisce per tutti i quattro tessuti, e non solo per Torino, valori della sigma della media che sono superiori a quelli calcolati col Metodo I. Se i sindonologi credono che ci sia stato un errore nel calcolo della sigma della media per Torino, devono credere che ci siano stati altrettanti errori anche per gli altri tre tessuti. Ma nessuno lo dice, tranne Brunati. Dopo il suo ultimo passaggio citato sopra (che si conclude con “è un vero e proprio dolo.”), Brunati prosegue [2]:

«Ed alla stessa conclusione si giunge considerando un altro fatto. Su tab. 2 di *Nature* si nota che non è solo sbagliata la tolleranza media del valore del campione 1, quello sindonico, ma anche quelle delle medie dei risultati ottenuti da quello stesso laboratorio per i campioni 2, 3 e 4, quelli di controllo. In tabella 2, infatti, i risultati sono i seguenti

[mostra i risultati per le quattro medie di Arizona dei quattro tessuti, e relative sigma, sia come forniti da *Nature* sia come calcolati da lui con le sigma più piccole.]

«Ma, anche qui, è evidente che debbono averlo fatto di proposito. Se, per il secondo campione, per esempio, la media delle età è stata di 927 anni, che è corretta, devono avere anche ottenuto, una

media delle tolleranze di 20 e non di 32. Altrimenti, ripeto, non sarebbero arrivati al 927. E lo stesso si può dire per il 1995  $\pm 46$  del campione 3 e per il 722  $\pm 43$  del 4.

«Salvo fare un'osservazione. Abbiamo visto come siano stati obbligati a sostituire il  $\pm 17$  del campione 1, quello sindonico, in  $\pm 31$  perché risultasse lecito il far la media finale. Un cambiamento scorretto, quindi, ma, almeno, giustificato. Non c'era nessun motivo, invece, che rendesse necessario trasformare i vari  $\pm 20$  in  $\pm 32$ ,  $\pm 46$  e  $\pm 43$ . Se, ripeto, la trasformazione del 17 in 31, poteva essere stata necessaria per far diventare pari al 5% quell'inaccettabile *significance level* pari al 1,04%, quale vantaggio si sarebbe ottenuto trasformando un *significance level* pari a 88%, come quello che si aveva se si adottava il  $\pm 20$  corretto, nel 90% che, invece, corrispondeva al  $\pm 32$ ? Sia l'88 che il 90% erano largamente superiori a quel 5% che costituiva il minimo. Quelle tre trasformazioni, quindi, devono essere state effettuate solo per confondere un po' le acque, per far sorgere in eventuali controllori il sospetto che i risultati ottenuti a Tucson richiedessero un trattamento diverso da quello richiesto da quelli di Oxford e Zurigo. Ulteriore imbroglio, quindi, che aggrava ulteriormente la posizione di quei signori.»

(Brunati considera accettabili i valori di  $p$  per gli altri tre tessuti, ma si riferisce al calcolo delle medie fra le medie dei tre laboratori. Invece vanno considerati i valori di  $p$  per le medie dei risultati all'interno di Arizona, e questi sono tutti inferiori o circa uguali a 0,05. Appunto per questo, Arizona ha dovuto allargare le sigma per tutti e quattro i tessuti.)

Non c'è bisogno di commenti. Secondo Brunati, gli autori di *Nature*, per nascondere di avere fraudolentemente alterato un dato, ne alterano altri tre!

Anche i sindonologi dovrebbero capire che se per tutti e quattro i tessuti Arizona presenta valori della sigma della media diversi da quelli calcolati da loro, non può trattarsi di quattro errori ma di un diverso metodo di calcolo. Dovrebbero anche accorgersi che la sigma risulta aumentata tanto più quanto è maggiore la dispersione dei dati, quindi seguendo una logica. Ma, a parte Brunati, nessun altro sindonologo, che io abbia notato, ne parla. In particolare Tosatti cita lunghi passaggi dall'articolo del 2005 di Brunati ma evita di parlare di questo punto.

Brunati torna a ribadire nell'articolo del 2006 [3]:

«Per analogia con i casi precedenti, se la media delle età era di 646 anni, la media delle tolleranze [intende sempre la tolleranza della media] doveva essere uguale alla radice quadrata dell'inverso di 0,00344595 essere, cioè, pari a  $\pm 17$ .

«Mentre, invece, in tab 2 risultava essere  $\pm 31$ . Un risultato in piena contraddizione con quel 646 che veniva indicato quale media delle età. Una notevole diversità che, al momento, non riuscivo proprio a spiegare.

«Per maggior sicurezza, ho sottoposto a verifica anche le 18 medie effettuate sui risultati (età e tolleranze) ottenuti dai tre laboratori quando avevano datato i tre campioni di controllo. 15 di queste medie, le 12 che riguardavano età e tolleranze rilevate dai laboratori di Zurigo ed Oxford e le tre riguardanti le misure delle età di Tucson erano corrette: il che confermava, fra l'altro, la validità del metodo di verifica adottato. Le tre medie delle tolleranze riguardanti le misure fatte a Tucson sui campioni di controllo, erano invece tutte errate.

«Sulla base delle mie verifiche avrebbero dovuto essere tutte uguali a  $\pm 20$ , mentre, secondo tab 2, risultavano pari a  $\pm 32$ ,  $\pm 46$  e  $\pm 43$ . Come mai? Tanto più che, anche in questi casi, il valore della media delle età era stato raggiunto passando da quel numero che avrebbe consentito di ottenere la media corretta delle tolleranze. Per esempio, la media delle età, 927 anni, fatta da Tucson sulla tela di Nubia, il campione di controllo nr 2, era stata ottenuta con una frazione il cui divisore era il numero 0,00344595, la radice quadrata dell'inverso del quale dava luogo a quel  $\pm 20$ , media corretta delle tolleranze. Come mai, allora, quel  $\pm 32$  di tab 2? E come mai, poi, per il campione di controllo 3 il  $\pm 20$  diventa  $\pm 46$ , e  $\pm 43$  per il campione 4? Pare evidente abbiano cambiato di proposito quei risultati per cercare di confondere le idee in caso di eventuali verifiche, per far insorgere il dubbio che, chissà perché, le misure fatte a Tucson richiedessero degli speciali sistemi per il calcolo delle medie. Un tentativo per giustificare il cambiamento di quel  $\pm 17$  in  $\pm 31$ ?

«Non ci potevano essere più dubbi, quindi, che quella fosse una indiscutibile alterazione del risultato.»

Brunati costruisce qui un castello così inverosimile che gli altri sindonologi non si azzardano a seguirlo, però tacciono sul fatto che per tutti i quattro tessuti di Arizona compare un analogo “errore”.

*§30. I sindonologi credono che, se il test del chi-quadrato non viene superato, i risultati siano da buttare.*

Se il test del chi-quadrato non viene superato, l'interpretazione ovvia nella nostra situazione è che le sigma dichiarate sono sottostimate, cioè non sono abbastanza grandi per spiegare le differenze fra i risultati delle varie datazioni. Non c'è alcun motivo per ritenere che le datazioni siano invalide o da rifare. L'unica conseguenza è che l'intervallo di fiducia per la media sarà più ampio di come sarebbe se lo si calcolasse dalle sigma dichiarate.

Non c'è niente di assoluto o definitivo nelle sigma dichiarate. Il laboratorio, limitatamente a quello che può quantificare, dà una stima provvisoria della precisione, ma non ha la certezza che la precisione sia proprio quella. Se, dopo i risultati di ripetute datazioni sullo stesso reperto, ci si accorge che le precedenti stime erano troppo basse, è naturale che si riformulino stime più adeguate. Se il test del chi-quadrato non viene superato, non si accetta la sigma della media quale calcolata con le sigma dichiarate e, come abbiamo visto nella prima parte, si cerca un altro metodo per calcolare l'ampiezza dell'intervallo di fiducia.

Non meraviglia affatto che un laboratorio, negli anni '80, dichiarasse valori sottostimati per le sigma. La sigma dichiarata descrive (o descriveva all'epoca) solo la variabilità fra le diverse misurazioni sulla stessa pasticca nella stessa seduta. Non tiene conto di quello che può succedere nelle fasi preliminari di preparazione del campione, prima dell'ultimissima fase del lancio degli ioni nel percorso dello spettrometro. Non tiene conto delle differenze che possono esserci fra due diverse sedute, tanto meno delle differenze fra due diversi laboratori. Anche per una stessa seduta, ci possono essere irregolarità non quantificate dal metodo che veniva usato per stimare la sigma dichiarata.

Che gli errori nelle datazioni possano essere maggiori di quanto previsto dalle sigma dichiarate, non è una ipotesi *ad hoc* che facciamo qui, ma è una constatazione generale per l'attività dei laboratori del radiocarbonio in quel periodo.

Quando, a partire appunto dagli anni '80, furono avviati programmi di intercomparazione fra diversi laboratori, si trovò che le differenze erano spesso superiori a quanto prevedibile in base alle sigma dichiarate. Rimandiamo alla terza parte dell'articolo dove mostriamo che differenze come quelle riscontrate per la Sindone, o anche maggiori, erano comuni.

Nelle seguenti citazioni abbiamo aggiunto i corsivi per sottolineare la ingiustificata convinzione dei sindonologi.

Cominciamo con due citazioni dal libro di Tosatti:

«Abbiamo sottoposto i dati a disposizione – l'articolo di “Nature” - a due professori di Statistica dell'università di Roma La Sapienza, Livia De Giovanni e Pierluigi Conti, totalmente estranei all'universo della Sindone. [p. 16] I risultati sono ancora più devastanti per la validità dell'esame. Il verdetto: le datazioni della Sindone dei tre laboratori non sono omogenee e *quindi l'esame avrebbe dovuto essere rifatto*. In più, uno dei tre laboratori, quello di Tucson, Arizona, *non avrebbe potuto nemmeno presentare i suoi risultati* per essere messi in “pool” con quelli di Oxford e Zurigo.» [11, p. 15]

«Dal momento che il laboratorio di Tucson non raggiungeva neanche il 5% di significatività nei

suoi risultati interni, in realtà la media non avrebbe dovuto essere fatta. *I laboratori avrebbero onestamente dovuto ammettere che l'esame non era riuscito.*» [11, p. 29]

Tosatti nel suo blog sul sito della Stampa [13]:

«L'errore di calcolo aveva permesso ai laboratori di affermare che avevano raggiunto il 5 per cento (la soglia minima) per considerare i risultati significativi. Conti e De Giovanni hanno dimostrato (l'analisi completa è nel libro "Inchiesta sulla Sindone, Piemme) che in realtà il test ha raggiunto al massimo l'uno per cento. *E di conseguenza l'esame avrebbe dovuto essere rifatto. E non avrebbe potuto essere dichiarato come riuscito.*»

Tosatti, qui e in generale, si fida dei suoi due consulenti e viene completamente fuorviato. Tosatti nell'intervista sul sito di Zenit:

«La "tolleranza" di errore che i tre laboratori si erano dati era del 5%; e dai numeri di *Nature* sembrava che si fosse raggiunto proprio il minimo, il 5%. In realtà è stato raggiunto l'uno per cento. *L'esame avrebbe dovuto essere rifatto, ma i campioni ormai erano distrutti.*» [9]

Sempre Tosatti nello speciale della *Stampa* all'epoca dell'ultima ostensione:

«Poi la rivista "Nature" pubblicò i risultati, e qui cominciarono i problemi. Perché i tre laboratori, come è d'uso, stabilirono che il loro esame poteva essere considerato attendibile purché i risultati verificati statisticamente, non scendessero sotto una certa soglia: il 5 per cento. Le cifre pubblicate su "Nature" firmate e controfirmate (compreso il British Museum, che doveva agire da "controllore") indicavano che era stata raggiunta proprio la soglia minima: il 5 per cento. Ma qualche tempo dopo, un appassionato della Sindone, l'ingegner Ernesto Brunati, si prese la briga di rifare i conti; e scoprì un errore di calcolo, inspiegabile ma fondamentale.

«Infatti senza quell'errore la soglia minima per dichiarare l'attendibilità dell'esame non si raggiungeva. E un esperto belga, Remi Van Haelst, confermava l'errore. Ma sia Brunati che Van Haelst sono dei fan della Sindone. E anche se i numeri sono numeri, e non ci si discute, scrivendo una "Inchiesta sulla Sindone" ho sottoposto quelle cifre a due specialisti della materia indipendenti e non sospetti di "sindonofilia", due docenti di Matematica e Statistica dell'Università La Sapienza, Luigi Conti e Livia De Giovanni. La risposta (corredata da una relazione scritta) dava ragione ai "franchi tiratori" sindonologi. In base ai dati forniti dai laboratori e pubblicati su "Nature", e rivisti dai professori della Sapienza, emerge che il massimo ottenuto è dell'1 per cento, non il fatidico e necessario 5 per cento.» [12]

Fanti nella intervista a Tornielli sul *Giornale*:

«Partiamo dalla datazione. Che cosa ha scoperto?

«"Non è una scoperta solo mia. Rifacendo i calcoli, sulla base dei dati forniti dai tre laboratori che eseguirono l'esame al radiocarbonio nel 1988, ci si rende conto che è stato commesso un errore. L'attendibilità della datazione medioevale è pari soltanto all'1,2%. *Cioè assolutamente inattendibile.*

«Com'è potuto accadere un errore di questo genere?

«"Diversi studiosi hanno dimostrato già da tempo che i risultati della datazione, in base al test statistico cosiddetto di Pearson, hanno una probabilità superiore al 95% di non corrispondere a quelli della Sindone. È stato inserito nella formula un numero sbagliato, che ha falsato il risultato finale, un 31 è stato sostituito con un 17. Questo farebbe pensare persino a una manomissione finalizzata a ottenere il risultato desiderato".

«È un'accusa grave...

«"I numeri sono numeri. E sono inattendibili. Queste confutazioni sono scritte da scienziati. Forse c'è una lobby che teme la verità su quei calcoli, teme di doversi rimangiare il risultato sulla Sindone

di origine medioevale, quando tutto, invece, lascia pensare che sia molto più antica e che risalga al primo secolo”.» [10]

Fanti nella sua relazione a un convegno del 2007:

«La datazione al carbonio 14 del 1988 è stata eseguita in modo discutibile perché, secondo diversi studiosi, il campione prelevato non è rappresentativo dell'intero Lenzuolo e potrebbe derivare da un rammendo medievale; inoltre il rapporto pubblicato su *Nature* contiene gravi errori statistici che pongono ulteriori interrogativi sul risultato [Nota 4].»

«[Nota 4] Per esempio, in accordo con l'ing. Brunati, il dott. Ballabio e R. Van Heest, appartenenti a ShroudScience, il valore errato dell'errore assegnato alla datazione della Sindone dal laboratorio di Arizona (31 anziché 17) comporta un abbassamento del livello di significatività statistica del campione dal 5% all'1% con la conseguente necessità di ripetere la prova.» [7, p. 9]

§31. *I sindonologi non si accorgono che ogni volta che il test chi-quadrato non viene superato, gli autori di Nature ne tengono conto usando un metodo di calcolo opportuno.*

È quello che è successo, come sappiamo, per le medie entro laboratorio di Arizona, dove si è usato il Metodo II, e per la media delle medie dei tre laboratori, dove si è usato il Metodo III.

Nel primo caso il resoconto non spiega quale sia il metodo usato, ma appare evidente che le sigma sono state allargate.

Nel secondo caso il resoconto spiega la situazione e i sindonologi non hanno scusanti. Infatti gli autori di *Nature* notano che: “The spread of the measurements for sample 1 [Torino] is somewhat greater than would be expected from the errors quoted”. Cioè dicono che hanno condotto il test del chi quadrato e hanno trovato che “it is unlikely that the errors quoted by the laboratories for sample 1 fully reflect the overall scatter.” (Ricordo che gli “errors quoted” sono quelli che qui indichiamo come le “sigma dichiarate”.) Dicono insomma che è improbabile che le sigma dichiarate rendano conto della dispersione. Quindi decidono “to estimate the uncertainty in that mean from the scatter of results”, cioè di stimare l'incertezza della media (fra i tre laboratori) dalla dispersione dei risultati. Abbiamo visto come è stato eseguito il calcolo con il Metodo III (§22).

§32. *Brunati scrive alla rivista Radiocarbon ma non comprende il senso della risposta che riceve.*

Nel 2006 Brunati pubblicò un articolo con il titolo “La corrispondenza con 'Radiocarbon' sulla datazione della Sindone”. Riferisce di avere scritto al *managing editor* di *Radiocarbon*, Mark McClure, a proposito dell'errore di Arizona (il famoso 31 al posto del 17). Dalla risposta ricevuta, Brunati crede che McClure gli dia ragione, ma non si accorge del significato di un numero, un 29, citato da McClure. Brunati è molto succinto nel riferire quanto risposto da McClure, ma si può supporre che McClure gli spiegasse che le sigma dichiarate erano state allargate col moltiplicatore, come abbiamo visto con il Metodo II.

Brunati scrive:

«Non ci potevano essere più dubbi, quindi, che quella fosse una indiscutibile alterazione del risultato. E l'Editor Manager di *Radiocarbon*, risponndomi con un Email datato 7 dicembre ultimo scorso [2005], ha dovuto ammettere l'errore, scrivendo *I agree it is not exactly 31 years* (riconosco che non sono esattamente 31 anni). Era la prima volta che un carbonista, un carbonista per giunta importante, riconosca che il rapporto su *Nature*, per quando riguardava la datazione della Sindone non fosse corretto.



«Anche se, subito dopo, chiude l'Email facendo presente che tale errore, in ogni caso, doveva esser considerato minimo. Secondo il McClure, infatti, gli scienziati del laboratorio di Tucson, essendo liberi di adottare il processo di calcolo che preferivano, avevano preso in considerazione la deviazione standard che portava la media delle tolleranze a 29 e non ai 17 anni indicati su *Nature*. Per cui l'errore invece d'esser costituito dalla differenza fra 17 e 31 anni, si riduceva a quello, più modesto, fra 29 e 31.» [3]

Da dove spunta quel 29? Brunati nei suoi calcoli non trova mai un 29. Trova un 17 che secondo lui dovrebbe essere al posto del 31 di *Nature* per la sigma della media di Arizona. Ma abbiamo visto (§19, Tab. 12) che applicando il moltiplicatore si trova appunto un 29 (o 28,97 col nostro calcolo a due decimali). Non è ancora il 31, che a quanto sembra McClure non sa spiegare, ma noi abbiamo supposto che la differenza sia dovuta all'aggiunta del termine per l'incertezza nella misurazione del delta-C13. Bisognerebbe sapere qualcosa di più del testo di McClure, ma mi sembra molto probabile che egli avesse calcolato il moltiplicatore e avesse trovato quel 29. Brunati non si rende conto di che cosa possa significare quel numero e continua:

«Non potevo esser d'accordo e l'ho scritto subito a *Radiocarbon* nel mio secondo Email del 6 dicembre. Non potevo esser d'accordo, in quanto il conteggio non era stato eseguito, come loro sostenevano, dal laboratorio, ma dal British Museum di Londra, al quale Tucson, come Zurigo ed Oxford, aveva inviato, appena rilevati, i suoi risultati parziali. Ed a Londra avevano adottato un unico criterio di calcolo per tutti, facendo le medie, come risultava su *Nature*, senza prender mai in considerazione quella deviazione standard alla quale si richiamava *Radiocarbon*. Validò, quindi il  $\pm 17$  e non il  $\pm 29$ .» [3]

In realtà, a quanto si desume dal resoconto di *Nature*, il British Museum accettò, per le medie entro laboratorio, i calcoli quali presentati da ciascun laboratorio. Anche se i calcoli per la media entro ciascun laboratorio fossero stati eseguiti a Londra, era legittimo, anzi doveroso, allargare le sigma per i soli dati di Arizona e non per i dati degli altri laboratori dove il test del chi-quadrato veniva superato.

Si ha una possibile conferma, per la nostra interpretazione di quel 29, anche da un altro passaggio di Brunati. Egli aveva risposto a McClure, a quanto sembra, tirando in ballo la sua teoria secondo cui il dato era stato alterato per poter far credere che venisse superato il test del chi-quadrato.

«Di conseguenza, l'Editor Manager di *Radiocarbon*, Mike [Mark] McClure, quando l'11 dicembre per la seconda volta mi ha risposto, ha dovuto ammettere: *we agree that the chi square test gives a result greater than one* (siamo d'accordo che la prova col chi quadro dia un risultato appena superiore all'unità). Col che, dopo avere riconosciuto che la media delle tolleranze non era corretta, tornava ad ammettere che il risultato della prova del chi quadro era tale da non consentire l'esecuzione della media finale e di conseguenza, di arrivare a concludere regolarmente la datazione.» [3]

In realtà il chi-quadrato nel caso in esame è molto più grande di uno (è 8,67, vedi §19, Tab. 12), e comunque non ha senso confrontare il valore del chi-quadrato con l'unità. Probabilmente McClure si riferiva al rapporto fra il chi-quadrato e il numero di gradi di libertà, cioè al moltiplicatore da applicare alla varianza. Se questo rapporto supera abbondantemente l'unità, significa che va usato un moltiplicatore. (Brunati traduce "risultato appena superiore all'unità" e non si capisce perché abbia aggiunto la parola "appena" che nel testo inglese non c'è.)

Brunati si illude di avere trionfato e continua:

«Se *Radiocarbon* era d'accordo sul fatto che quella datazione non era stata una cosa seria, era però anche ben deciso a non pubblicarlo. Adducendo la scusa che, comunque, quella era una vecchia pratica a proposito della quale si era già tanto scritto e discusso.

«Mi pare che un simile atteggiamento sia più che mai discutibile. È vero che molto è stato scritto a quel proposito, ma lo si è sempre fatto basandosi su ipotesi piuttosto vaghe. Ora, invece, eravamo di fronte a conclusioni ben chiare alle quali si era giunti basandosi su motivazioni sicure, della quali loro stessi avevano riconosciuto la validità. Mi pareva che questo dovesse farli sentire obbligati a rendere pubblica la nuova situazione che si era determinata. Come, a suo tempo, avevano trionfalmente annunciato che quel lenzuolo, essendo medioevale, non poteva aver avvolto il Corpo di Cristo, oggi, a maggior ragione, dovevano sentire il dovere di confessare d'essersi sbagliati.»

### §33. *Tosatti è stato sfortunato nella scelta dei suoi consulenti di statistica.*

I due consulenti di Tosatti, i professori Livia De Giovanni e Pierluigi Conti, non gli hanno dato buoni consigli.

I due statistici hanno firmato una analisi riportata in appendice al libro di Tosatti [11, pp. 193-196]. Hanno eseguito calcoli per la verifica di omogeneità dei dati e hanno concluso che il test non è superato né per il confronto fra le medie dei tre laboratori per Torino né per i dati interni di Arizona. Avrebbero dovuto dire a Tosatti che questa non è una loro scoperta, ma che già gli autori del resoconto di *Nature* se ne erano accorti e avevano tenuto conto della disomogeneità. Poi avrebbero dovuto spiegare a Tosatti che se il test di omogeneità non è superato, ciò non significa che i risultati siano invalidati, ma significa solo che occorre tenerne conto nel calcolo di un intervallo di fiducia, ciò che è stato fatto dagli autori di *Nature*.

Inoltre i due professori nei calcoli per il confronto fra i tre laboratori (ultima tabella a p. 195) hanno usato il valore 17, invece di 31, per la sigma della media di Arizona. Quindi hanno accettato il valore calcolato da Brunati (e analogamente da Van Haelst, da loro citato come fonte), senza accorgersi che su *Nature* per il calcolo della sigma della media di Arizona era stato usato, e doveva essere usato, un altro metodo.

I due hanno anche dimostrato di non conoscere la situazione del radiocarbonio negli anni '80 e non hanno spiegato a Tosatti che simili variabilità nei risultati non erano eccezionali ma erano quanto ci si poteva aspettare.

Conti compare nel trailer di un documentario (*La notte della Sindone*. al momento non ancora in vendita) dove dice:

«I numeri che sono stati ottenuti, sono troppo diversi per essere spiegabili dall'errore di misura, e quindi significa che i laboratori stavano datando materiale intrinsecamente diverso. Cioè quello che stava datando un laboratorio non era esattamente coincidente con il materiale, la stoffa, che stava datando un altro laboratorio.»

Lo dice lui in prima persona, quindi non è stato Tosatti a fraintendere il suo pensiero.

Per la cronaca, Conti e la De Giovanni sono o sono stati anche docenti alla LUMSA, cioè alla Libera Università Maria Santissima Assunta.

In conclusione, spero che i sindonologi imparino qualcosa dalla lettura del presente testo. Non so se mi ringrazieranno.

### III – La qualità delle datazioni negli anni '80

I sindonologi sembrano meravigliarsi se nei dati per la datazione della Sindone ci sono notevoli differenze fra le diverse misurazioni effettuate nello stesso laboratorio o in diversi laboratori. Sembrano non capire che le sigma dichiarate (*quoted errors*) possono essere sottostimate e non rendere conto della effettiva variabilità dei risultati. Non si rendono conto che se nel confronto fra diverse datazioni dello stesso reperto non viene superato il test di omogeneità del chi-quadrato, questo può significare semplicemente che le sigma dichiarate sono sottostimate e devono essere allargate per essere conformi alla realtà. Diamo allora un quadro di quale fosse la situazione per il livello di precisione che veniva conseguito nei laboratori all'epoca della datazione della Sindone.

#### §34. Le sigma dichiarate possono essere sottostimate

Come abbiamo già visto (§2), ma conviene ripeterlo, per ciascuna pasticca immessa nello spettrometro il C14 viene misurato molte volte, diverse decine. Dalla variabilità dei risultati, il laboratorio stima la sigma. Questa variabilità tiene conto solo di ciò che accade dentro al percorso dello spettrometro per una singola pasticca, nella stessa seduta e a confronto con gli stessi standard e gli stessi bianchi.

Ma le incertezze nella datazione non dipendono solo da fattori che operano all'interno dello spettrometro. Dipendono anche da altri fattori, che non vengono considerati nella stima della sigma dichiarata. Prima di entrare nella macchina, il campione subisce diversi trattamenti. Viene applicata la procedura di pulizia, che può non asportare tutto l'inquinante o per converso può introdurre carbonio recente presente in laboratorio. Poi c'è la combustione, poi la trasformazione in grafite, un processo complesso che richiede diverse fasi. Poi c'è la ionizzazione. Durante tutte queste fasi possono essere introdotti fattori di incertezza che *non* vengono considerati nel calcolo della sigma dichiarata. Un errore dell'1% nella misurazione del C14 comporta uno spostamento nella datazione di oltre 80 anni. Quindi è normale che confrontando risultati ottenuti con diversi campioni provenienti dallo stesso reperto, come i dodici campioni della Sindone, possa risultare una variabilità un po' superiore rispetto a quella teoricamente prevedibile dai valori delle sigma dichiarate. I dodici sottocampioni della Sindone hanno avuto storie separate a partire da quando sono stati sottoposti a pulizia (effettuata a volte con metodi diversi). Inoltre le operazioni sono state svolte in tre laboratori diversi.

Perché allora i laboratori forniscono un valore “dichiarato” di sigma che può essere inferiore a quello reale? Per il semplice motivo che è l'unico che sono in grado di quantificare. Ciascuna pasticca viene misurata molte volte, ma di solito non succede che da un singolo reperto vengano fatti molti campioni e vengano tutti datati ripetendo le procedure fin dall'inizio. È successo con la Sindone che venissero datati dodici frammenti, e allora è normale che fra i dodici ci sia una variabilità superiore a quella che c'è fra le diverse misurazioni dello stesso frammento.

Insomma la sigma dichiarata dice solo quale fosse la precisione della macchina durante quella particolare seduta. Non tiene conto di tutto quello che è successo prima che la pasticca venisse introdotta nella macchina.

Non solo. Anche considerando soltanto quello che avviene all'interno dello spettrometro durante una seduta, si può riscontrare una variabilità superiore al previsto. Se ne ha qualche esempio appunto nei dati della Sindone e dei tre tessuti di controllo. Arizona, come sappiamo (§20), ha fornito datazioni separate per coppie di pasticche ottenute dallo stesso sottocampione e presumibilmente trattate assieme nelle fasi preliminari, e anche fra due pasticche gemelle si sono talvolta riscontrate notevoli differenze..

Inoltre nei dati di Arizona per gli altri tessuti, come si può vedere sul resoconto di Nature, sono riportati cinque risultati, benché i campioni fossero sempre divisi in quattro, perché di uno dei sottocampioni, evidentemente in due pasticche, vengono forniti due risultati, che però questa volta

non sono accorpati e sono tenuti distinti e dichiarati. (Probabilmente Arizona intendeva fare lo stesso anche per Torino, fornendo otto risultati e dichiarando che a due a due erano ottenuti dallo stesso sottocampione. Secondo le informazioni raccolte da Remi Van Haelst, fu il British Museum a chiedere ad Arizona di accorpare le coppie di risultati per Torino.)

Avendo a disposizione i dati per coppie di pasticche ottenute con lo stesso processo dallo stesso sottocampione e misurate assieme, è interessante notare come ci possano essere forti discrepanze anche fra i due risultati di una coppia. Come si vede nella Tab. 13 di §20, per le quattro coppie di Torino c'è una differenza di 32, 121, 136, 0. Per gli altri tre tessuti, nelle pasticche gemelle si trovano differenze di 167, 58, 162. Si vede che, su sette confronti, si trovano in quattro casi differenze superiori a 120 anni, e per due di questi le differenze sono superiori a 160 anni. Questo non si spiega con le sigma dichiarate che sono attorno a valori di 40 o 50. Si vede quindi che a maggior ragione le sigma dichiarate tendono a essere sottostimate e che fattori di errore intervengono non solo durante le fasi preliminari, ma anche fra pasticche "gemelle" ottenute allo stesso modo e misurate simultaneamente.

Va anche detto che i risultati di Arizona entro ciascun tessuto sono nettamente quelli che mostrano maggiore variabilità, rispetto agli altri laboratori, facendo supporre che la precisione fosse minore per Arizona. La misurazione del radiocarbonio nello spettrometro è un processo complesso e delicato e se non si riesce a tenere tutti i parametri sotto controllo, emerge qualche irregolarità nei risultati. Ricordiamo ancora che un piccolo errore dell'uno per cento nella misurazione della quantità di C14 produce uno scarto di 83 anni nell'età.

Nel resoconto è anche indicato che Oxford trovò un risultato anomalo per una delle misurazioni ripetute durante una stessa seduta. Lo scartò come "outlier", anche se apportò una modifica al risultato. Quindi non solo con due pasticche gemelle, ma con la stessa pasticca e durante la stessa seduta, si possono verificare forti discrepanze. Infatti per una delle datazioni di Oxford per il tessuto della Nubia, è stata apposta questa nota:

«One anomalous replicate (of 6) obtained for independent measurement O2.2b; if rejected it reduces date by 40 yr; final date quoted actually reduced by 20 yr.»

Sembra di capire che si fecero sei datazioni sulla stessa pasticca (presumibilmente in sei giri di tamburo) e una di queste diede un risultato di una età tanto più giovane che eliminandolo la media si riduceva di 40 anni. La media denunciata è stata ridotta di 20 anni. La media denunciata per quel sottocampione è stata di 925 anni. Significa che il risultato anomalo era di oltre 200 anni più giovane. Eppure la sigma dichiarata per quel sottocampione è solo 45 anni.

A partire dagli anni '80, sono stati saltuariamente condotti programmi di confronto fra diversi laboratori. Chi organizzava il confronto prendeva molti campioni dello stesso materiale e li spediva a diversi laboratori, ciascuno dei quali eseguiva la datazione. Facendo il confronto, si è visto che le differenze fra le diverse datazioni tendevano a essere maggiori rispetto a quanto prevedibile dalle sigma dichiarate.

Dopo che un laboratorio aveva disponibili dati di confronto, diventava possibile fare una statistica e vedere quanto, mediamente e per quel laboratorio, la sigma dovrebbe essere aumentata per tener conto dei reali risultati. Qualche laboratorio ha introdotto l'uso di un fattore moltiplicativo, per esempio 1,3 o 1,5, da applicare alla sigma dichiarata. Però il fattore moltiplicativo è solo ipotetico e non si sa quale dovrebbe essere in ciascun caso particolare. Quindi la soluzione migliore è lasciare la sigma dichiarata, quale ricavata dalla variabilità interna alla macchina in ciascuna seduta, e avvisare il cliente che quello è un valore minimo e che l'incertezza potrebbe in realtà essere maggiore. Si trova per esempio che la statunitense Beta Analytic, azienda leader nei servizi di datazione, in un suo foglio di informazioni per i clienti scriveva nel 1999:

«The quoted errors represent 1 sigma statistics. Since these errors cannot include uncertainties outside of those which can be quantified during measurement, it is best to consider them as minimum quotes.»

Questo mi sembra un modo appropriato per presentare il risultato di una datazione.

### §35. Il confronto fra sei laboratori

Nel 1983-84 fu condotta una prova che doveva servire a dimostrare che era possibile datare la Sindone con i nuovi metodi per piccoli campioni. I sei laboratori che si erano candidati per la datazione della Sindone si accordarono per misurare in cieco campioni provenienti dagli stessi tessuti. A fornire i campioni e a coordinare il lavoro fu nominato Michael Tite del British Museum di Londra. Appunto in seguito a questo suo coinvolgimento, Tite sarà anche il coordinatore per la datazione della Sindone nel 1988. Il resoconto fu pubblicato nel 1986 a firma dello stesso Tite e di altri due autori del British Museum, Richard Burleigh e Morven Leese. La Leese sarà anche fra gli autori del resoconto su *Nature* per la Sindone.

Fra i laboratori, quattro usavano l'allora nuovissimo metodo spettrometrico con acceleratore (AMS): Arizona, Oxford, Zurigo e Rochester. Di questi, i primi tre sono gli stessi che poi faranno la datazione della Sindone. Il laboratorio di Zurigo non era ancora pronto per eseguire il pretrattamento dei campioni che fu eseguito presso il laboratorio del radiocarbonio di Berna. Anche Rochester fece eseguire la preparazione a un altro laboratorio a Reston in Virginia. Poi c'erano due laboratori col metodo dei minicontatori, una evoluzione per piccoli campioni del tradizionale metodo del conteggio dei decadimenti radioattivi: Harwell (Inghilterra) e Brookhaven (New York). Nel resoconto viene conservato l'anonimato dei laboratori, cioè vengono forniti i risultati ma senza specificare quale laboratorio abbia ottenuto ciascun risultato. La precauzione serviva a evitare che qualche laboratorio si facesse cattiva fama se forniva risultati scadenti.

Era previsto un confronto per due antichi tessuti, poi ne venne aggiunto un terzo. Il primo era un tessuto egiziano di lino, gli altri due erano tessuti peruviani di cotone.

Qualche laboratorio datò più volte lo stesso tessuto, qualche altro non datò tutti i tessuti, da qui le disparità nel numero dei risultati per ciascun tessuto.

Ci furono due risultati, uno per il primo e uno per il terzo tessuto, che erano in forte disaccordo rispetto agli altri. Nel resoconto non viene fornito il nome del laboratorio responsabile. Si venne poi a sapere che la colpa era del laboratorio di Berna nella fase di pretrattamento dei campioni per Zurigo. Il laboratorio aveva usato un nuovo metodo per la pulizia dei campioni. Non viene specificato in che cosa consistesse questo nuovo metodo e doveva essere un metodo strano perché in un caso si ottenne una data più giovane di 860 anni, nell'altro più vecchia di 960 anni rispetto alla media degli altri laboratori. Dopo che furono constatate le discrepanze, vennero forniti a Berna due nuovi frammenti per ciascuno dei due tessuti. Fu eseguito il trattamento di pulizia con un metodo tradizionale e questa volta Zurigo fornì date in accordo con gli altri laboratori. L'episodio venne ignorato come un incidente. Le due date incongrue non vennero considerate nell'analisi dei dati e non le considereremo nemmeno qui. Includeremo invece i valori ottenuti da Zurigo nelle prove ripetute dopo che Berna ebbe rimediato all'errore nel pretrattamento.

Veniamo ai risultati. Consideriamo dapprima la precisione raggiunta, cioè le differenze fra le varie datazioni di uno stesso tessuto. La Tabella 15 ha tre colonne per i tre tessuti usati per il confronto, rispettivamente Egitto, Perù 1 e Perù 2. Abbiamo aggiunto una colonna per i risultati poi ottenuti nel 1988 per la Sindone.

Tabella 15

	Egitto	Perù 1	Perù 2	Sindone
Numero datazioni	9	5	6	12
Dispersione	147	178	94	61
Media sigma dichiarate	109	98	107	48
Differenza estremi	437	450	240	204

Nella prima riga c'è il numero delle datazioni, che come abbiamo detto non è lo stesso per tutti i tessuti.

La seconda riga mostra la dispersione dei risultati. Come si vede, i valori della dispersione furono molto maggiori che per la Sindone.

La terza riga mostra la media delle sigma dichiarate che per i tre tessuti è molto maggiore che per la Sindone. Il test del chi-quadrato, calcolato con le sigma dichiarate, non viene superato per Perù 1 e supera appena il 5% per Egitto. Va considerato che con valori così grandi per le sigma dichiarate, il test viene superato più facilmente.

La quarta riga mostra la differenza fra i valori estremi (in anni BP). Anche qui i risultati per la Sindone sono i migliori con una differenza fra estremi che è più piccola rispetto ai tre tessuti. È da notare che il numero di risultati per i tre tessuti è minore che per la Sindone (9, 5 e 6 rispetto a 12), quindi la differenza fra gli estremi aumenterebbe, probabilmente, se si aggiungessero altri risultati. Se per esempio scelgo a caso 5 risultati fra i 12 della Sindone, avrei solo una probabilità piuttosto piccola (15%) di scegliere anche i due estremi.

In conclusione la precisione è stata molto migliore per la Sindone rispetto ai tre tessuti usati nella prova di confronto.

Passiamo a considerare l'accuratezza. Per valutarla dovremmo conoscere, indipendentemente da queste misurazioni, la vera età di ciascun reperto. Purtroppo l'età non era nota da considerazioni archeologiche sicure e per i tessuti peruviani, in particolare, era stimata soltanto dallo stile e con una incertezza di diversi secoli.

Il primo tessuto era un lino egiziano che per la sua provenienza si riteneva databile a circa il 3000 a.C. I risultati, benché con una dispersione notevole, nell'insieme abbracciano quel periodo. La media delle nove datazioni è 4257 BP corrispondente a una data calibrata del 2895 a.C. L'intervallo di fiducia è 3255-2827 a.C.

Il secondo tessuto, che abbiamo indicato come Perù 1, fornì una età che era parecchio più giovane di quanto ci si attendesse. La media dei risultati portava alla prima metà del Cinquecento. Scartando un risultato con età nettamente più giovane, si andrebbe alla metà del Quattrocento. L'età era prevista al 1200 circa, in base allo stile che indicava la cultura Chimù. Non so quanto fosse sicura quella stima del 1200. Guardando su internet trovo che la cultura Chimù sembra estendersi da circa il 1000 fino al 1470 quando il territorio fu conquistato dagli Inca. Se la data del 1200 fu presa semplicemente come media fra questi estremi, e non in base a specifici dettagli stilistici, non si potrebbe escludere che la reale età del tessuto fosse del Quattrocento inoltrato. (Del resto mi chiedo se non sia possibile che la gente del posto continuasse ancora per qualche tempo a tessere con lo stile tradizionale anche dopo il 1470). Non risulta che in seguito siano state tentate altre datazioni su quel tessuto. Insomma non sappiamo se la datazione dei laboratori fosse nell'insieme sbagliata e forse non lo era. Comunque fu preferito continuare il confronto fra i laboratori con un altro tessuto, anche questo peruviano, ma anche per questo non era nota con precisione l'età.

Per il secondo cotone peruviano, Perù 2, lo stile indicava un "tardo periodo intermedio" e la data prevista era nell'arco 1000-1400, un'estensione molto ampia che non permetterebbe di sapere se i

risultati siano in buon accordo con la reale età. La media dei risultati dei laboratori porta al 1400. Si può notare che Tite non fu molto efficiente nel procurare tessuti di età nota. Per i due tessuti peruviani, la stima dell'età fu fatta soltanto in base a considerazioni stilistiche perché non se ne conosceva la provenienza archeologica.

Vediamo quindi che i risultati dei sei laboratori furono assai meno buoni di quelli poi ottenuti con la Sindone. È da notare che anche così il risultato fu giudicato soddisfacente, quindi i laboratori non si aspettavano di meglio. Infatti nella conclusione del rapporto del 1986 si legge:

«Overall, there is good agreement between the results obtained and the expected historical dating of the samples, in particular as far as Samples 1 and 3 are concerned.

«A coherent series of results can be obtained when several laboratories undertake separate blindfold measurements of the same sample.» [4]

E infatti i laboratori si ritennero confermati nella loro decisione di candidarsi a datare la Sindone. È vero che questo succedeva nel 1983-84 quando i metodi per piccoli campioni erano ancora nuovi e stavano migliorando, ma infatti nel 1988 i risultati per la Sindone furono nettamente migliori e non ci si poteva aspettare di più.

### §36. Un programma di intercomparazione.

All'epoca della datazione della Sindone, era in corso il primo vasto programma di intercomparazione fra diversi laboratori, organizzato con base a Glasgow in Scozia. Partecipavano molti laboratori su scala internazionale. I laboratori con AMS che portarono a termine il programma furono solo cinque. A tutti i laboratori venivano fornite porzioni degli stessi reperti o materiali. I laboratori eseguivano le datazioni in condizione di cieco e comunicavano i risultati agli organizzatori che poi eseguivano i confronti. I risultati furono pubblicati nel 1990 in diversi articoli riuniti in un numero di *Radiocarbon* (Volume 32/3).

Il quadro che emerse mostrò che le discrepanze fra i laboratori erano superiori rispetto a quanto ottimisticamente ci si aspettava.

In un articolo riassuntivo di uno degli organizzatori, Murdoch Baxter, si legge [1]:

«Do quoted errors account for the observed variability?»

«Generally, no. Two laboratories grossly overestimate errors but most labs seriously underestimate their errors, by a factor of 2 to 3 times. Only 7 labs from 38 passed all three very basic desirable performance criteria at each stage of the study.»

Baxter si riferiva a laboratori che erano in maggioranza quelli coi metodi tradizionali dei conteggi dei decadimenti, ma non è che all'epoca si ritenesse che il metodo AMS dovesse essere migliore, se mai il contrario.

Ricordo che i “quoted errors” sono per noi le sigma dichiarate. Nella frase citata, Baxter si chiede se le sigma dichiarate rendano conto della variabilità osservata. La risposta è no. La maggioranza dei laboratori sottostimava le sigma per un fattore anche di due o tre volte. (Quindi i sindonologi non dovrebbero meravigliarsi se per la Sindone in qualche caso le sigma fornite dai laboratori erano leggermente sottostimate.)

Per avere un'idea delle reazioni nell'ambiente, possiamo rifarci a una intervista allo stesso Baxter che fu pubblicata nel 1989 sulla rivista *New Scientist* in occasione di un congresso in Scozia dove furono discussi i risultati. Baxter era direttore di un laboratorio del C14 in Scozia ed era uno degli organizzatori del programma di intercomparazione.

L'articolo di Andy Coghlan [5] ha il titolo: “Unexpected errors affect dating techniques “ e comincia così:

«The margin of error with radiocarbon dating, an analytical method for finding out the age of ancient artefacts, may be two to three times as great as practitioners of the technique have claimed. The shortcomings of the method, revealed earlier this month at a workshop at East Kilbride near Glasgow in Scotland, mean that while some laboratories consistently date artefacts correctly almost to the year, others are up to 250 years out.»

Poi continua:

«Britain's Science and Engineering Research Council (SERC) commissioned a trial that compared the accuracy with which 38 laboratories around the world dated artefacts of known age. Of the 38, only seven produced results that the organisers of the trial considered to be satisfactory.

Murdoch Baxter, the director of the Scottish Universities Research and Reactor Centre at East Kilbride near Glasgow, and one of the organizers of the trial, said that the survey represented “a major turning point in the history of the method”. His laboratory was one of only three from Britain that participated [...]

«All 38 participants had to date a set of samples made from wood, peat and carbonate. The laboratories involved were on average “two to three times less accurate than implied by the range of error they stated”, said Baxter.»

I laboratori usavano in maggioranza uno dei due metodi convenzionali del conteggio dei decadimenti radioattivi, ma anche in quel caso le sigma dichiarate erano calcolate principalmente in base alla statistica dei conteggi e non tenevano conto delle fasi preliminari di preparazione. L'articolo prosegue:

«Most of the errors quoted [le sigma dichiarate] acknowledge uncertainties in the accuracy with which the pulses of radioactivity from the sample are counted. “It is now clear,” says Baxter, “that other unaccounted-for sources of error occur during the processing and analysis of samples.” He suspects that the most serious unforeseen errors arise in the chemical pre-treatment of samples.»

Poi confronta i diversi metodi:

«Baxter says that accelerator mass spectrometry, used last year by a laboratory at the University of Oxford to date the Shroud of Turin, allegedly the burial shroud of Jesus Christ, came out of the survey badly. Five of the 38 participating laboratories used this technique, for which samples weighing a few milligrams are acceptable. The other techniques require grams of the sample. Baxter said that some of the accelerator laboratories were way out when dating samples as little as 200 years old.

«Because so little material is used in accelerator mass spectrometry, the effects of chemical pre-treatment are likely to be more serious, says Baxter. “The samples are probably more prone to atmospheric dust or dandruff,” he said.»

Baxter, che non disponeva di AMS, rifletteva forse la rivalità che c'era in quegli anni fra i laboratori col metodo tradizionale e i laboratori col nuovo metodo AMS. Forse trattava i laboratori AMS peggio di come meritassero. In realtà guardando i dati sugli articoli di *Radiocarbon*, mi sembra che i risultati del metodo AMS non siano peggiori degli altri. Ci furono laboratori convenzionali che fecero di peggio. Però su cinque laboratori AMS, ce ne furono due che mostrarono un errore sistematico, uno in un senso e l'altro in senso opposto, attorno ai cento anni.

I laboratori che partecipavano erano anonimi, cioè i risultati venivano pubblicati senza precisare i nomi dei laboratori. Nell'intervista Baxter non si lascia sfuggire l'occasione di dire che il laboratorio AMS di Oxford non aveva partecipato. Trattandosi di una iniziativa presa in Gran Bretagna, il fatto che Oxford si astenesse dal partecipare al confronto potrebbe far sospettare che temesse di non figurare benissimo. In quegli anni era anche fresca la polemica (di cui Baxter non parla) per i



risultati discordanti che erano stati ottenuti nelle ripetute datazioni di un famoso reperto, il cosiddetto “Uomo di Lindow”: Il laboratorio di Oxford e quello convenzionale di Harwell avevano ripetutamente datato il reperto e Oxford forniva costantemente una età più vecchia di circa trecento anni.

L'articolo conclude:

«In the light of the results, researchers are to adopt new practices to improve quality control. One is to increase the frequency with which laboratories have samples “blind” checked by others. Another, backed by the International Atomic Energy Agency, is to improve the standard reference materials of known age for analysis to test their machines' accuracy. The agency will distribute a standard set of materials from 1990.»

Infatti negli anni seguenti fu prestata attenzione a migliorare il controllo di qualità, che fino ad allora era stato carente come dimostravano i risultati dell'intercomparazione.

### §37. Le sigma dichiarate per la Sindone erano particolarmente piccole

Il fatto che fosse necessario un moltiplicatore  $\sigma$ , ciò che è lo stesso, che il test di omogeneità non venisse superato, dipendeva dai valori delle sigma dichiarate. Ho quindi controllato se i valori di sigma dichiarati per la Sindone e per gli altri tre tessuti del 1988 fossero in linea con quelli prevalenti all'epoca.

Ho confrontato con gli elenchi di tutte le datazioni pubblicate da Oxford nel periodo 1988-1991. Ho considerato solo le età non superiori a 2000 anni. Ho trovato 173 datazioni (escludendo quelle della Sindone e dei tre controlli che appartengono esse pure all'elenco) che hanno una media delle sigma dichiarate pari a 76.

Ho guardato anche un elenco di Harwell con metodo convenzionale pubblicato nel 1990, considerando le età non più vecchie di 3000 anni. La media delle sigma dichiarate è 77.

Queste medie di 76 o 77 si riferiscono alle sigma quali dichiarate e non aumentate con moltiplicatori o altro.

Per confronto, la media delle sigma dichiarate fu 55 per la Sindone e 52,5 includendo anche gli altri tre tessuti di controllo.

Quindi quella volta le sigma dichiarate furono particolarmente piccole e non ci si meraviglia se non rendevano conto della variabilità dei risultati.

In anni recenti, negli ultimi elenchi pubblicati, Oxford arriva a valori di sigma dichiarata attorno a 30 anni o anche meno.

### §38. Se si ripetesse oggi la datazione

Dal 1988 a oggi molte cose sono cambiate. Il controllo di qualità è migliorato, almeno per i laboratori più qualificati. Con la prassi instaurata negli anni successivi, ogni laboratorio nel corso del tempo ripete datazioni dello stesso materiale. Per controllare non solo la precisione ma l'accuratezza, esegue regolarmente datazioni di materiale di età nota e partecipa all'occasione a confronti con altri laboratori.

Si è prestata attenzione a procurarsi materiale per gli standard che sia davvero standard, ciò che non è semplice con l'aumentare del numero dei laboratori e della quantità delle datazioni.

I parametri di funzionamento dell'apparato vengono costantemente monitorati e se insorge qualche difetto il lavoro viene interrotto e si cerca di porre rimedio.

Simili precauzioni cominciavano a essere adottate negli anni '80 ma non in modo così sistematico.

Dopo il primo vasto progetto di intercomparazione, che come abbiamo visto si concluse nel 1989, sono stati attuati altri programmi simili con partecipazione di un maggior numero di laboratori. Il

numero di laboratori AMS partecipanti è salito a diverse decine. Si è constatato che i laboratori AMS hanno raggiunto un buon grado di affidabilità. Se raramente emerge qualche risultato anomalo, bisogna considerare che durante gli stessi anni il numero dei laboratori è aumentato e quindi ci sono laboratori in funzione da poco che non hanno ancora messo a punto il lavoro. Non così affidabili risultano i laboratori convenzionali, in particolare quelli nella versione a scintillatore liquido. Per questi si sono verificati casi di errori anche di diversi secoli. Ma non significa che i metodi convenzionali siano peggiori. Infatti i laboratori che producono seri errori sono in numero limitato e alcuni di loro non hanno messo a punto le procedure, per esempio non hanno un buon controllo sulla qualità dello standard moderno che impiegano.

Negli anni recenti i migliori laboratori, nella datazione di campioni di età nota (non troppo antica), arrivano spesso a ridurre lo scarto a meno di cinquanta anni, fino anche a circa trenta anni. Però non mancano gli scarti superiori a cinquanta anni e talvolta, in casi sfortunati, superiori a cento. Insomma se la datazione della Sindone venisse ripetuta oggi, si potrebbe sperare in un risultato più preciso, ma anche un risultato come quello del 1988 non sfigurerebbe.

## Bibliografia

- [1] M.S. Baxter: Report of the International Workshop of Intercomparison of Radiocarbon Laboratories: *Radiocarbon* 32/3, 1990, 389-391  
Nello stesso fascicolo di *Radiocarbon* ci sono diversi articoli con i risultati di quel confronto. Ulteriori analisi sono nel volume 34/3 del 1992. Per un rapporto su un programma di confronto più recente, vedere per esempio, sempre su *Radiocarbon*, il volume 45/2 del 2003.
- [2] E. Brunati: Altro che rammendi! La datazione della Sindone è tutta un falso. *Collegamento pro Sindone Internet*, Maggio 2005  
<http://www.sindone.info/BRUNATI1.PDF>
- [3] E. Brunati: La corrispondenza con “Radiocarbon” sulla datazione della Sindone. *Collegamento pro Sindone Internet*, Febbraio 2006)  
<http://www.sindone.info/BRUNATI2.PDF>
- [4] R. Burleigh, M. Leese, M. Tite: An intercomparison of some AMS and small gas counter laboratories. *Radiocarbon* 28/2A, 1986, 571-77
- [5] A. Coghlan: Unexpected errors affect dating techniques. *New Scientist*, 30 Settembre 1989
- [6][3] P. E. Damon et al: Radiocarbon Dating of the Shroud of Turin. *Nature*, 337. 16 febbraio 1989, pp. 611-615  
<http://www.shroud.com/nature.htm>
- [7] G. Fanti: La Sindone: un mistero che sfida la Scienza. (Relazione a un convegno a Lorenzago di Cadore, 19 luglio 2007.)  
<http://www.dim.unipd.it/fanti/Lorenzago.pdf>
- [8] G. Fanti: *La Sindone, Una sfida alla scienza moderna*. Aracne, 2009
- [9] A. Gaspari: Chi è l'uomo della Sindone? Il vaticanista Marco Tosatti racconta l'inchiesta sul telo più visitato al mondo. (6 aprile 2009)  
<http://www.zenit.org/article-17836?l=italian>
- [10] A. Tornielli: «Il Medioevo? Data inattendibile. Il lenzuolo risale al primo secolo» *Il Giornale*, 19 marzo 2008.
- [11] M. Tosatti, *Inchiesta sulla Sindone*, Piemme, 2009
- [12] M. Tosatti: Sul Sacro Telo l'ombra del complotto. Le analisi scientifiche rivelano sorprendenti imprecisioni: perchè? *La Stampa* (speciale ostensione), 9 aprile 2010
- [13] M. Tosatti: Sindone e C14: esame fallito, il perché. (Blog sul sito della *Stampa*, 11 aprile 2010.)
- [14] R. Van Haelst: Radiocarbon Dating The Shroud. A Critical Statistical Analysis.  
<http://www.shroud.com/vanhels3.htm>
- [15] G.K. Ward e S.R. Wilson: Procedures for Comparing and Combining Radiocarbon Age-Determinations: A Critique. *Archaeometry*, 20 (1), Febbraio 1978. pp. 19-31.

[16] [http://en.wikipedia.org/wiki/Weighted\\_mean](http://en.wikipedia.org/wiki/Weighted_mean)

[17] <http://www.c14dating.com/frac.html>