Application of statistical shape analysis to the classification of renal tumours appearing in childhood

M.A./Dipl. Math. /Cand. Soz. päd.
Stefan Markus Giebel
Stefan.Giebel@t-online.de
Université du Luxembourg Campus Limpertsberg
Dr. J.P.-Schenk Uniklinikum Heidelberg
Prof. Dr. J. Schiltz, Uni Luxembourg

Overview

1) Survey

2) Shape

3) Mean Shape

4) Tests

5) Differences between tumours shown in an optical way

6) Classification
- Conclusion
- Forecast

Renal tumours appearing in early childhood

- Wilms- tumours growing in the near to the kidney
  Genetic cause
  The majority of renal tumours in the childhood is diagnosed as "Wilms" (80%)
  There are four types of tissue (a, b, c, d) and three stages of development (I, II, III)

- Renal cell carcinoma
  growing also in the near to the kidney
  Are very rare in the childhood

- Neuroblastoma
  Growing in the near to nerve tissue
  Also very rare

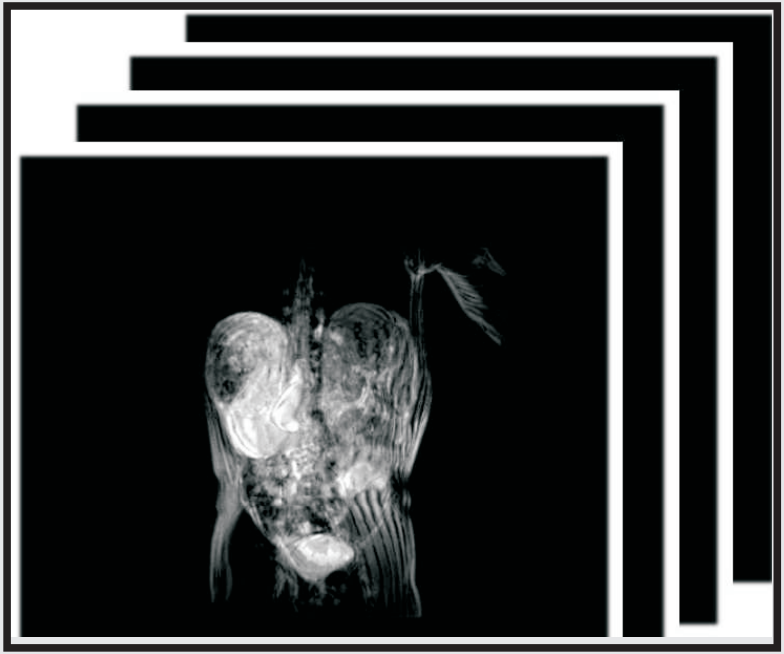
- Clear cell carcinoma
  Growing in the near to the bones .
  Are also rare

- etc.

The therapy depends on the diagnosis.

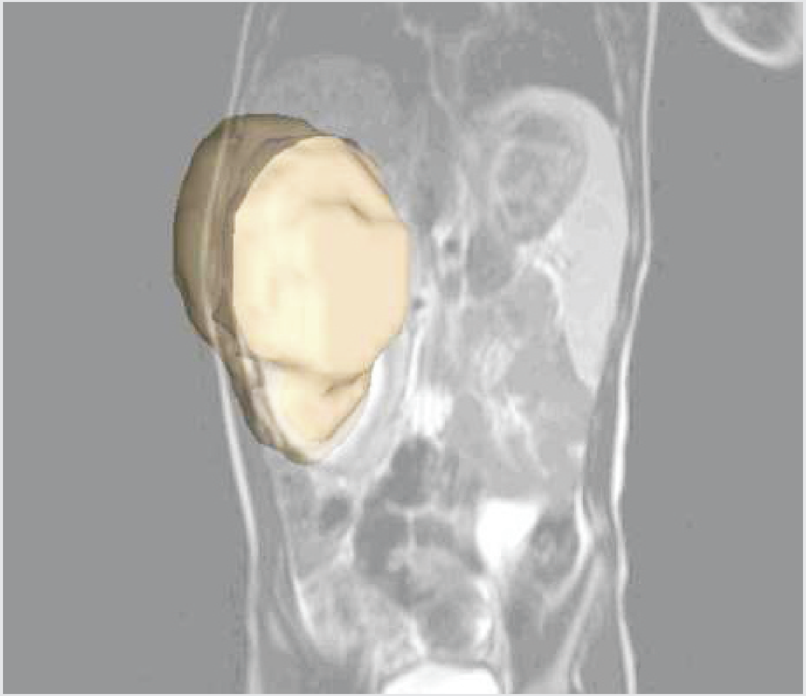
Reference: "Die Mischgeschwülste der Niere" (Wilms, 1899)

Dicom Data



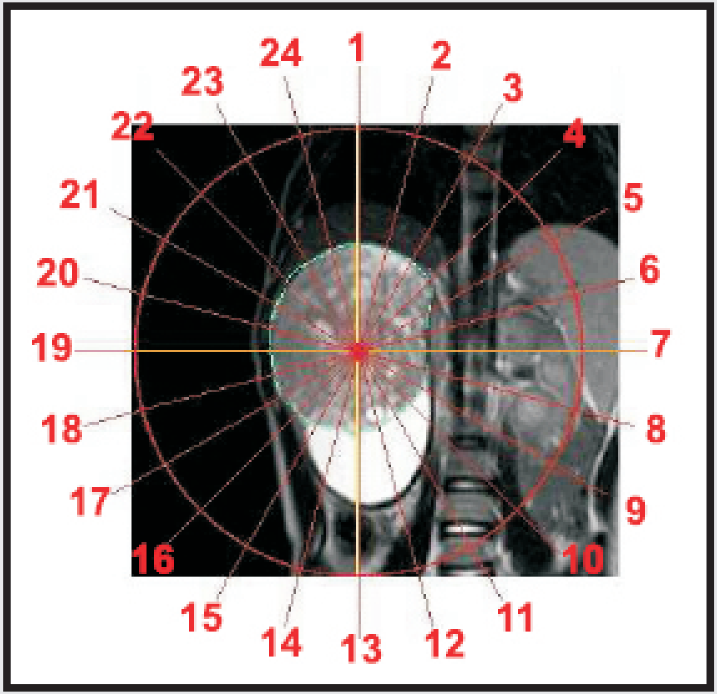
Getting transversal and frontal images of the tumour
Problem: Not for each patient we have both views
Images created by Magnetic resonance tomography

Three-dimensional object



1. Construction by using the data (density, depth etc.)

Explorative* survey of landmarks

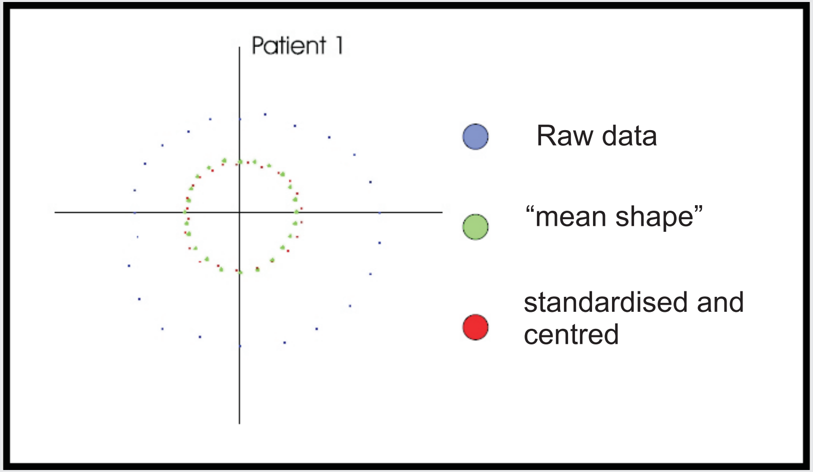


*there are no medical relevant points used as landmarks

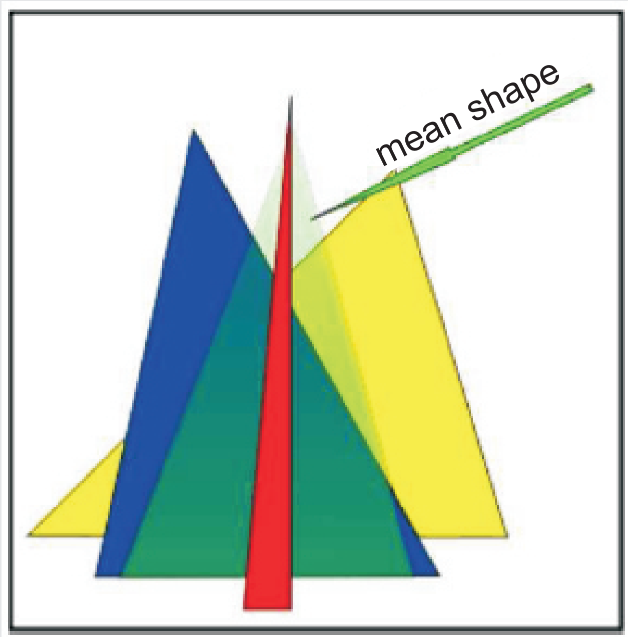
1. Determining of three dimensional mass point
2. Taking two dimensional image therein the mass point

Data process

1. Standardisation (using Euclidean norm)
2. Centring on two-dimensional centre

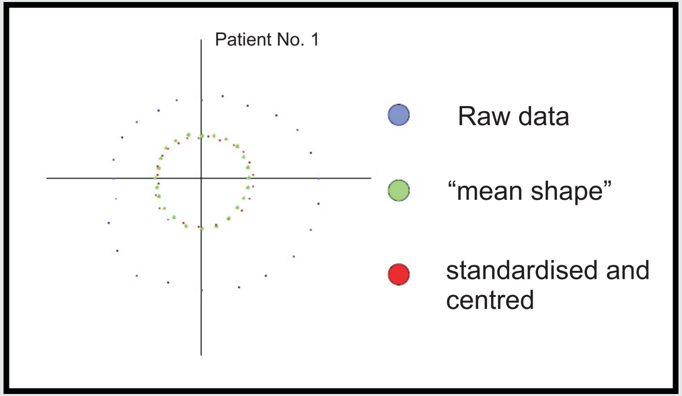


Determining of „mean shape“



Determining the expected “mean shape” of a group of objects.
That mean's: smallest distance in the average to all shapes in the group

Determining of “mean shape”



Using the following algorithm

$$i = 1, \dots, n$$

$$\bar{m} \mapsto w_i(\bar{m}) = \begin{cases} \frac{\langle \bar{m}, o_i \rangle}{|\langle \bar{m}, o_i \rangle|} & \text{if} \quad \langle \bar{m}, o_i \rangle \neq 0 \\ 1 & \text{if} \quad \langle \bar{m}, o_i \rangle = 0 \end{cases}$$

$$\bar{m} \mapsto T(\bar{m}) = \frac{1}{n} \sum_{i=1}^{n} w_i(\bar{m}) o_i$$

recursively there is a sequence

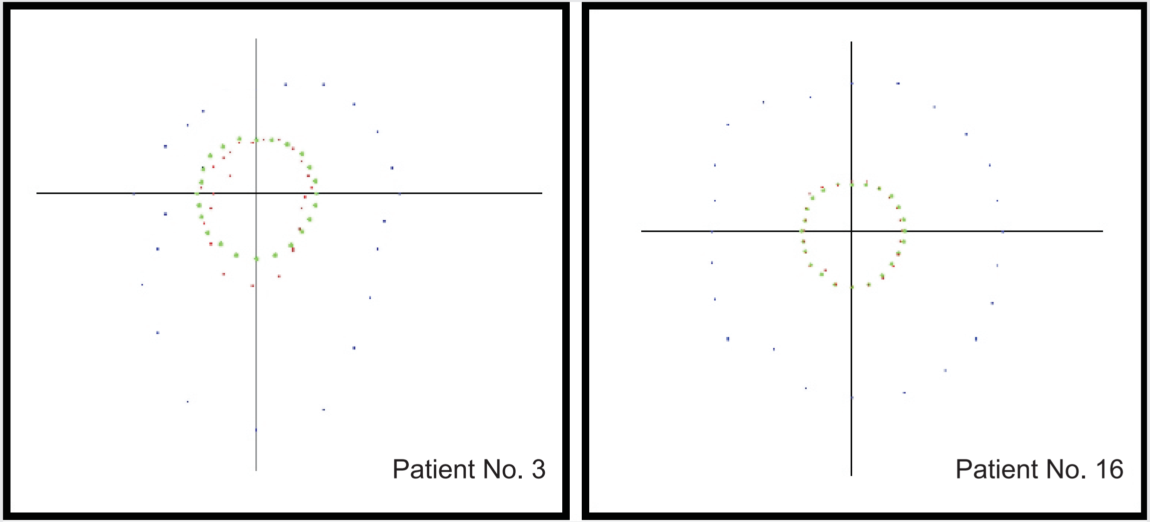
$\bar{m}_r = T(\bar{m}_{r-1}), r = 1, 2, \dots$ iterations

criterion to stop

$$\bar{m} = T(\bar{m})$$

Algorithm for „mean shape“ (Ziezold 1994)

Determining of „mean shape“



Statement: Patient No. 3 is very far from the “mean shape”.
Patient No. 16 is very near to the “mean shape“.

Distance from the "mean shape" (Wilms)

patient		distance	
Nr.	Diagnose	d_f	$rang_W$
Nr.1	n.b.	0.0849	3
Nr.2	IId	0.1009	6
Nr.3	IIc	0.2260	18
Nr.4	IIIa	0.0968	5
Nr.5	IIa	0.1567	13
Nr.6	IIb	0.1113	8
Nr.7	IId	0.1940	17
Nr.8	IId	0.1448	12
Nr.9	IId	0.1854	16
Nr.10	IIc	0.1290	11
Nr.11	IIb	0.1834	15
Nr.12	IIa	0.0772	2
Nr.13	IIc	0.0916	4
Nr.14	IIc	0.1058	7
Nr.15	IIc	0.1126	9
Nr.16	n.b.	0.0541	1
Nr.17	IIa	0.1178	10
Nr.18	IIc	0.1754	14

Description of the test (Ziezold, 1994)

The group of m objects is an indepent realisation of the distribution P and the other group of k objects an independent realisation of the distribution Q

Determining of p-value according to the test
H_o P=Q

$H_{1:}$ $P \neq Q$

1. step: Determining of “mean shape”

2.step: Detemining of distances to the “mean shape” and the u_0 according to the Mann Whitney-U-Test

3.step: Determining all possible u-values separating the group (k+m) in two groups with m and k objects

4.step: Determining the rank of u_0 in the group of all u-values
5.step: p-value = r/N

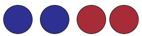
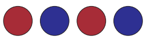
6.step: Determining the p-values in the other direction. Determining “mean shape” in the group of m objects

Description of the test (Ziezold 1994)

High u_O-values means: A lot of cases - not used for the “mean shape”- has a smaller distance to the “mean shape” than the cases used for the “mean shape”

Low u_O-values means: Only a small number of cases - not used for the “mean shape”- has a smaller distance to the “mean shape” than the cases used for the “mean shape”

Determining of all possible permutations
possibilities

bbrr 
rrbb ....
rbbr ....
brrb ....
brbr ....
rbrb 

$4!/(2!\,2!) = 6$ possibilities

$|All|/(|subset_1|!\,|subset_2|!$ = Number of all possibilities

Checking of differences between types of „Wilms“- tumours

Subsets		Differentiation					
Tumortyp 1	Tumortyp 2	u_0	$m_=$	$m_<$	$p - Intervall$	k	$\binom{15}{k}$
$Typ\ a$	$\overline{Typ\ a}$	0	57	0	$[0.002, 0.125]$	3	455
$\overline{Typ\ a}$	$Typ\ a$	21	14	338	$[0.745, 0.774]$	12	455
$Typ\ b$	$\overline{Typ\ b}$	2	22	64	$[0.619, 0.819]$	2	105
$\overline{Typ\ b}$	$Typ\ b$	9	5	37	$[0.362, 0.409]$	13	105
$Typ\ c$	$\overline{Typ\ c}$	6	17	431	$[0.086, 0.090]$	6	5005
$\overline{Typ\ c}$	$Typ\ c$	14	155	780	$[0.156, 0.187]$	9	5005
$Typ\ d$	$\overline{Typ\ d}$	17	52	970	$[0.711, 0.749]$	4	1365
$\overline{Typ\ d}$	$Typ\ d$	10	40	153	$[0.113, 0.141]$	11	1365

m= …: Number of cases with the same u-value
m<…: Number of cases with a lower u-value
The interval is a result of the smallest and the highest rank of u_o

Checking of differences between different tumours

N1:neuroblastoma
N2: renal cell carcinoma
K: clear cell carcinoma

Subsets		Differentiation						
Tumortyp 1	Tumortyp 2	u_0	$m_=$	$m_<$	$p - Intervall$	k	n	$\binom{n}{k}$
Wilms	N1	12	47	122	[0.0924, 0.1271]	3	21	1330
N1	Wilms	15	36	834	[0.6271, 0.6541]	18	21	1330
Wilms	K	5	4	13	[0.0737, 0.0895]	2	20	190
K	Wilms	0	103	0	[0.0053, 0.5421]	18	20	190
Wilms	N2	11	3	11	[0.6667, 0.7778]	18	19	18
K	N1	0	7	0	[0.1, 0.7]	2	5	10
N1	K	1	2	5	[0.6, 0.7]	3	5	10
K	N2	0	3	0	[0.3333, 1]	2	3	3
N1	N2	1	2	1	[0.5, 0.75]	3	4	4

m= …: Number of cases with the same u-value
m<…: Number of cases with a lower u-value
The interval is a result of the smallest and the highest rank of u_o

Conclusions

„Typ c“ and clear cell carcinoma have a tendency for differentiation

Neuroblastoma only in one direction

Renal cell carcinoma not differentiable

Independence (Influence) of Landmarks of Shapes

Ziezold •Mathematische Schriften Kassel, Heft 03/2003

H_0 The k th landmark of X is independent (influenced by) of the other landmarks with respect to the distance

H_1 The k th landmark of X is **not** independent (influenced by) of the other landmarks with respect to the distance

Step 1.

$$A_p = A_p^1 = \frac{1}{n - n_p} \sum_{d(\mathbf{x}^{(i)\prime}, \mathbf{x}^{(j)\prime}) > c_p} \frac{|x_{ik} - x_{jk}|}{d(\mathbf{x}^{(i)\prime}, \mathbf{x}^{(j)\prime})} .$$

Distance between Landmarks

Distance between Objects without kth landmark

$$A_p^s = \frac{1}{n - n_p} \sum_{d(\mathbf{x}^{(i)\prime}, \mathbf{x}^{(j)\prime}) > c_p} \frac{|x_{\tau_s(i),k} - x_{\tau_s(j),k}|}{d(\mathbf{x}^{(i)\prime}, \mathbf{x}^{(j)\prime})}$$

Random selected

N: 100 possibilities

n: all cases
n p: all possibilities for 2 in n cases
p: only a part of the sample (p-quantille)

Step 2.

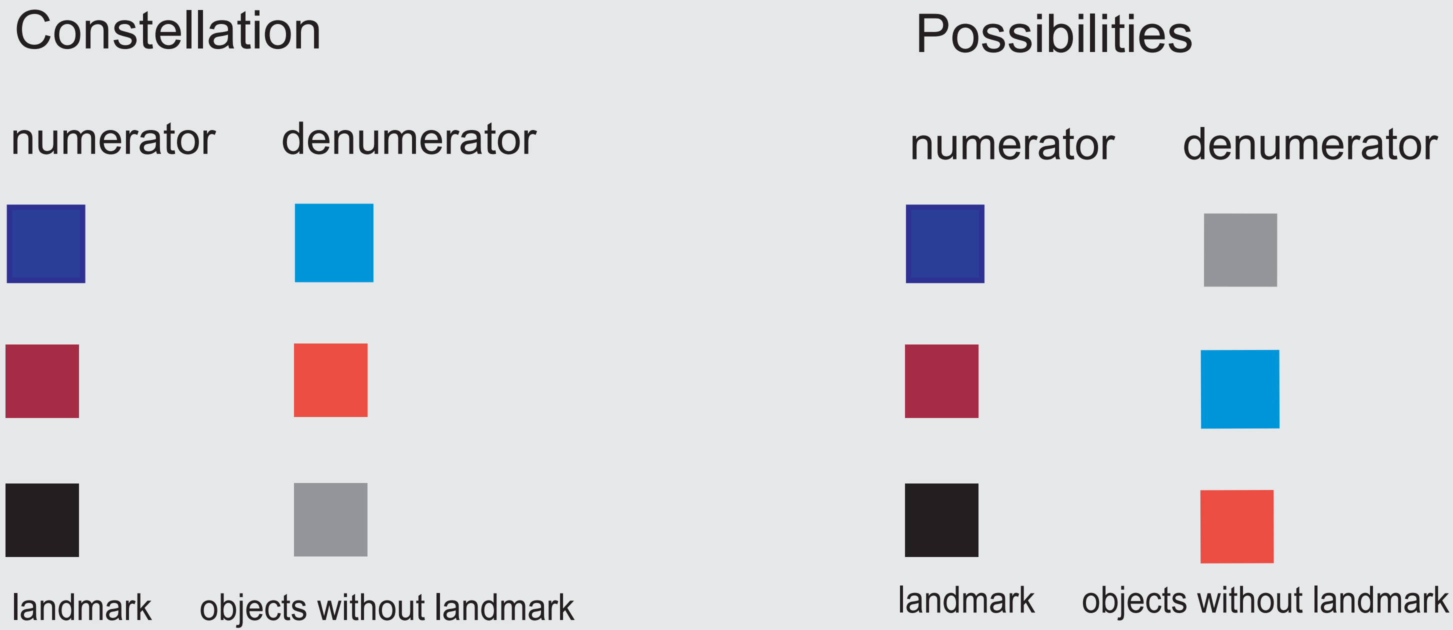
$$R_p = \operatorname{rank}(A_p^1, \{A_p^1, A_p^2, \ldots, A_p^N\})$$

Step 3.

$$\pi_p = \frac{N - (R_p - 1)}{N} \leq \alpha \text{ i.e. } R_p \geq N(1 - \alpha) + 1$$

p - value

Explanation of test

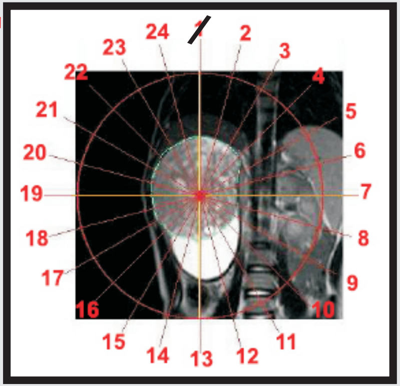


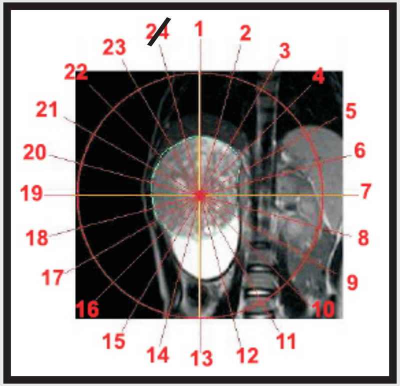
Results
for
p-quantile= 65%

N=100

all landmarks are independent of the other landmarks with respect to the distance

Landmark	Rank
1	45/46
2	48/49
3	34/35
4	39
5	54/55
6	43
7	40
8	39
9	33
10	46/47
11	48
12	45/46
13	43
14	35/36/37
15	42
16	34
17	30
18	30
19	32
20	35
21	31/32/33
22	49
23	43
24	42





Wilcoxon - Test

Also it is interesting to test the distance of landmarks to the mean shape for differentiating nephroblastoma to neuroblastoma. For that test we use the Wilcoxon-Test and calculate according to the Mann-Whitney-U-Test all possibilities. We assume that the <u>average of difference</u> to the mean shape for every landmark can be used for differentiating the tumors.

Wilms / Neuroblastome

$p \in [0,1887; 0,1917]$

Neuroblastome / Wilms

$p \in [0,7586; 0,763]$

No results for $\propto = 0,1$

Wilms Landmark $\xrightarrow{d}$ Mean shape Landmark

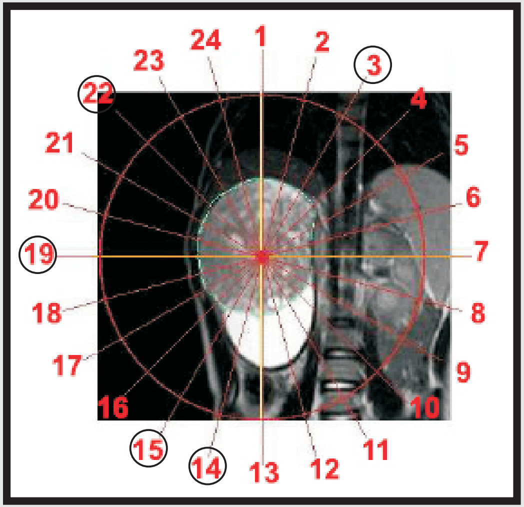
Neuro Landmark $\xrightarrow{d}$ Mean shape Landmark

<u>The average is not enough for differentiation</u>

Explorative take k=5 landmarks from 24

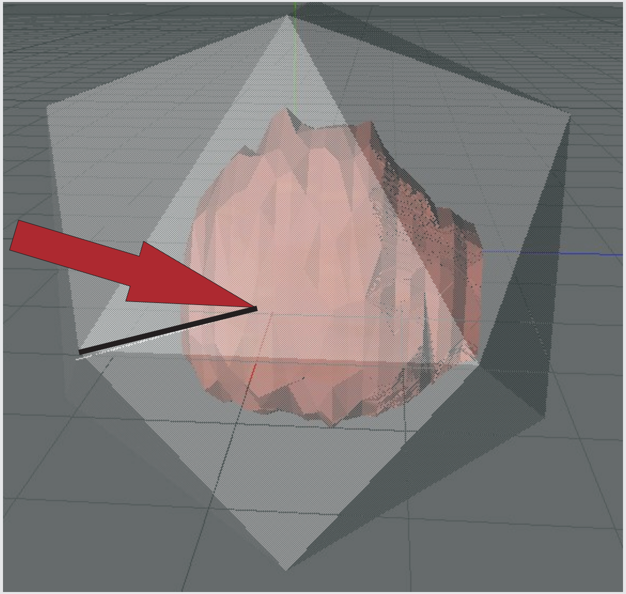
1. One sample for best configuration (Test Ziezold 1994)
   (smallest u-value for differentiating neuro/wilms and wilms/neuro)

2. One sample for test the configuration



Forecast

Determining of three dimensional landmarks



First results

Three dimensional case

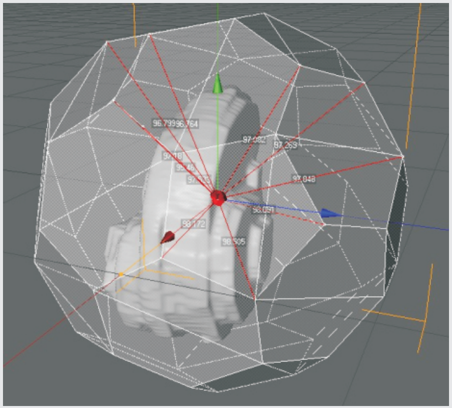
Sample:
5 neuroblastoma - 14 wilms
Test Ziezold (1994)

Wilms / Neuroblastoma

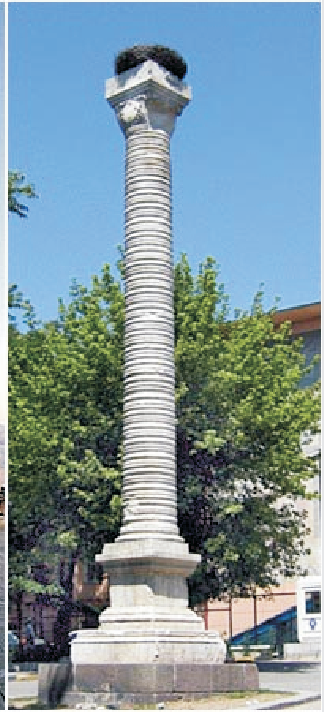
$p \in [0{,}157; 0{,}187]$

Neuroblastoma / Wilms

$p \in [0{,}069; 0{,}108]$



15 from 60 landmarks

Teşekkür ederim

Application of statistical shape analysis to the classification of renal tumours appearing in early childhood

(M.A./Dipl. Math. /Cand. Soz. päd.)
Stefan Markus Giebel
Stefan.Giebel@t-online.de
Université du Luxembourg Campus Limpertsberg
Dr. J.P.-Schenk Uniklinikum Heidelberg
Prof. Dr. J. Schiltz, Uni Luxembourg