

3 Ergebnisse

3.1 CT-Messungen

Auf allen CT-Aufnahmen, unabhängig von der benutzten Schichtdicke, waren die Zirbeldrüsen gut zu erkennen. Sowohl ihre Größe als auch ihr Verkalkungsgrad zeigen in der untersuchten Stichprobe erhebliche Variationen. Die Abgrenzung des Drüsengewebes vom umgebenden Wasser bereitet keine Probleme. Ebenfalls unproblematisch ist die Abgrenzung des Stäbchens, an dem die Drüse befestigt ist, vom oberen Drüsenende. Bei keiner der hier untersuchten Drüsen waren Zysten zu erkennen. Manche Drüsen zeigen solitäre Verkalkungen, die scharf begrenzt und gegenüber dem umgebenden Gewebe stark hyperdens sind. Größere Konkremeente können sogar isodens mit dem im Phantom eingebauten Hydroxylapatit-Zylinder sein. Es gibt auch Drüsen, die diffuse, weniger stark hyperdense Verkalkungen enthalten. Mit Hilfe der 1mm- und der 2mm-Methode kann man Form und Größe der Kalkablagerungen besonders gut erkennen. Aber auch wenn die benutzte Schichtdicke 4 mm oder 8 mm beträgt, ist der verkalkte Anteil des Drüsengewebes gut zu beurteilen.

Volumenmessung

Die Volumenmessung bildet die Basis für die Beurteilung des funktionellen Zustandes der Zirbeldrüse. Das Volumen der Drüse zeigt erhebliche interindividuelle Unterschiede. Die Volumenbestimmung alleine sagt aber nichts über die Melatoninproduktion aus. Dennoch müssen wir neben dem Verkalkungsgrad auch das Drüsenvolumen kennen, um genauere Angaben über den funktionsfähigen Drüsenteil machen zu können.

Bei der Volumenbestimmung der Drüsen mittels CT (1mm-Schichten) wurden Werte zwischen 38 mm^3 und 286 mm^3 ermittelt. Bei der ersten Messung betrug der niedrigste Wert 47 mm^3 , der höchste Wert lag bei 273 mm^3 . Bei der zweiten Messung lag das Minimum bei 38 mm^3 , das Maximum bei 286 mm^3 . Nach der Methode von Bland und Altman wurde zu jeder Drüse die Differenz aus den beiden Meßwerten gebildet. Bei der 1mm-Messung betrug die mittlere Differenz zwischen den beiden Volumenwerten $M2 - M1$ 7 mm^3 . Die Varianz der einzelnen Wertepaare, die für die statistische Auswertung entscheidend ist, ergab Werte zwischen $0,5$ und 450 mm^6 . Die mittlere Wertepaar-Varianz war $41,5 \text{ mm}^6$.

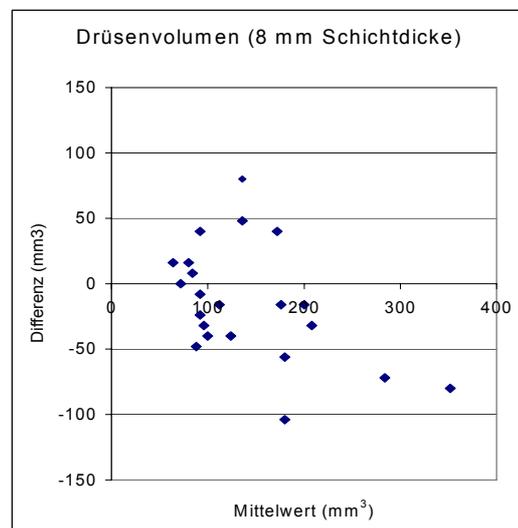
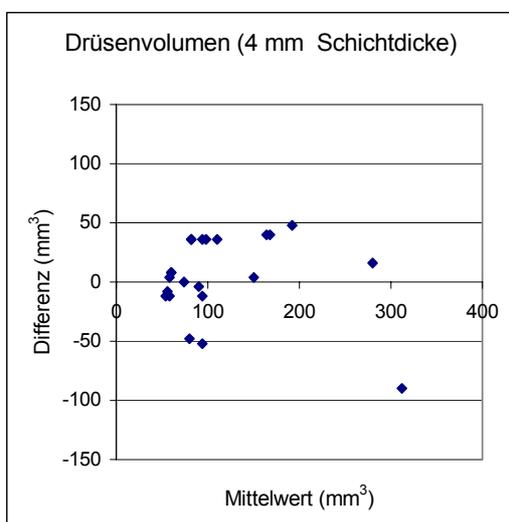
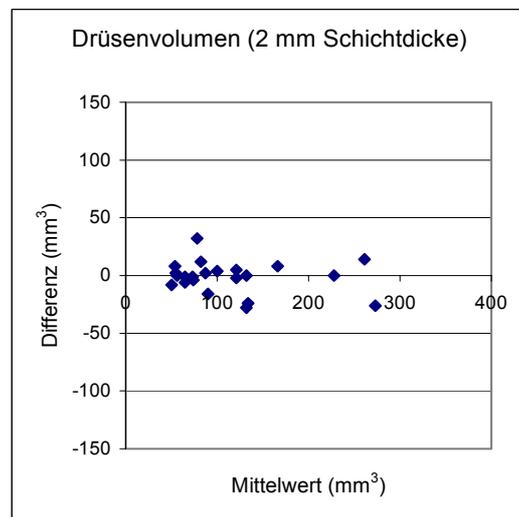
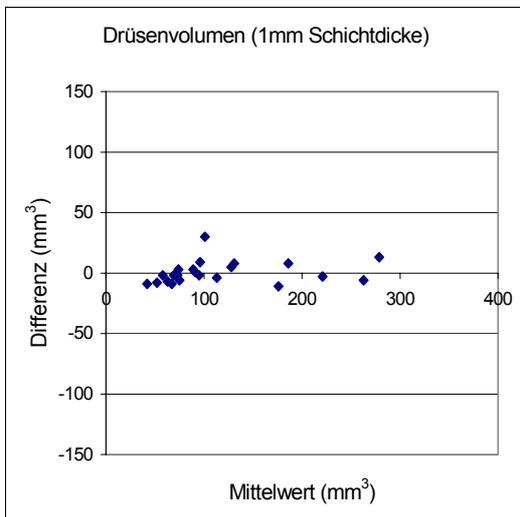
Bei der 2mm-Messung wurden Volumenwerte zwischen 46 mm^3 und 286 mm^3 registriert. Bei der ersten Messung schwankten die Werte zwischen 50 mm^3 und 286 mm^3 und bei der Wiederholungsmessung zwischen 46 mm^3 und 268 mm^3 . Die mittlere Differenz zwischen Erstmessung und Zweitmessung betrug in diesem Fall 12 mm^3 . Die Varianz der Wertepaare schwankte zwischen 0 mm^6 und 1250 mm^6 . Der Mittelwert der Varianz war 152 mm^6 .

Bei den 4mm-Schichten wurden Ergebnisse zwischen 48 mm^3 und 356 mm^3 gemessen. Bei der ersten Messung ergaben sich Werte zwischen 56 mm^3 und 356 mm^3 . Bei der zweiten Messung waren die Ergebnisse im Bereich zwischen 48 mm^3 und 288 mm^3 . Die mittlere Differenz war 27 mm^3 . Die Wertepaarvarianz ergab Werte zwischen 8 mm^6 und 3872 mm^6 . Die mittlere Wertepaarvarianz ist 584 mm^6 .

Bei der 8 mm-Messung liegen die Werte zwischen 64 mm^3 und 392 mm^3 . Die erste Messung ergab Meßwerte zwischen 72 mm^3 und 392 mm^3 . Bei der zweiten Messung waren es zwischen 64 mm^3 und 312 mm^3 . Die mittlere Differenz beträgt 40 mm^3 . Die Wertepaarvarianzen bewegen sich zwischen 0 mm^6 und 5408 mm^6 , der Mittelwert ist 1238 mm^6 .

Diagrammübersicht 1

Reproduzierbarkeit der CT-Volumenmessung in Abhängigkeit von der Schichtdicke



In diesen Diagrammen sind die Differenzen der beiden Volumenmessungen gegen den Mittelwert aufgetragen. Diese graphische Darstellung folgt dem von Bland et al. entwickelten Verfahren. Es ist deutlich zu erkennen, daß die Differenzen mit zunehmender Schichtdicke auch größer werden. Das deutet auf eine bessere Reproduzierbarkeit bei Nutzung dünnerer Schichten hin.

UCPT-Messung

UCPT ist in unserem Modell eigentlich die entscheidende Größe. UCPT steht für uncalcified pineal tissue und korreliert signifikant mit der Melatoninproduktion des Organismus. Eine Verringerung würde auf einen Rückgang des melatoninproduzierenden Anteils der Zirbeldrüse hinweisen. Deswegen dürfte diese Größe bei späteren klinischen Untersuchungen von zentraler Bedeutung sein. Genau wie CPT ist UCPT eine willkürlich festgelegte Größe und wird deswegen in VU (volume units) und nicht in mm^3 angegeben. Die statistische Auswertung der UCPT-Messung erfolgte nach dem gleichen Prinzip wie die Auswertung der Volumen- oder CPT-Messung.

Die 1mm-Messung ergab UCPT-Werte zwischen 0 VU und 191 VU. Bei der Erstmessung wurden Werte zwischen 0VU und 191 VU registriert; bei der Wiederholungsmessung schwankten die Werte zwischen 0 VU und 188 VU. Die mittlere Differenz zwischen Erstmessung und Wiederholungsmessung wurde zu 1,8VU errechnet. Die für den späteren statistischen Vergleich entscheidende Wertepaarvarianz liegt zwischen 0VU^2 und 312VU^2 . Die mittlere Wertepaarvarianz beträgt 44VU^2 .

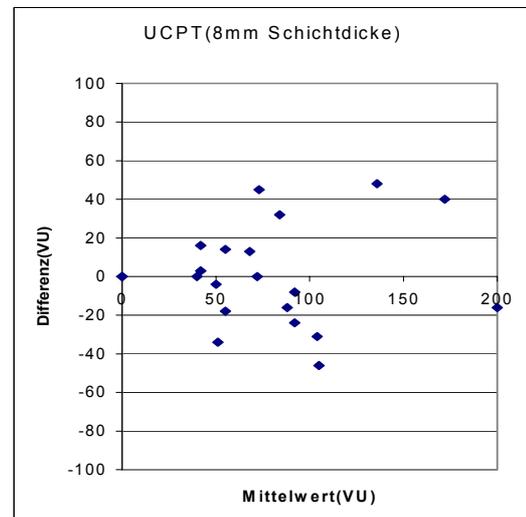
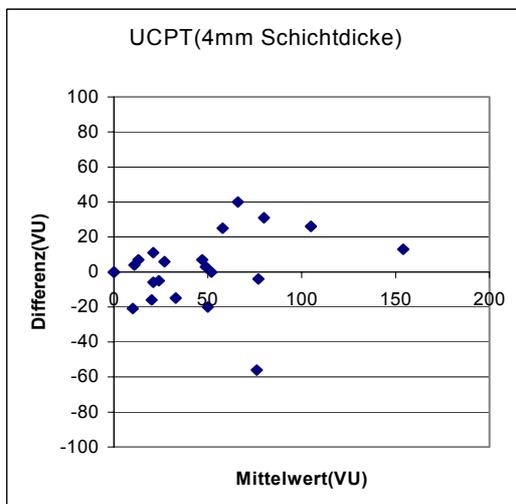
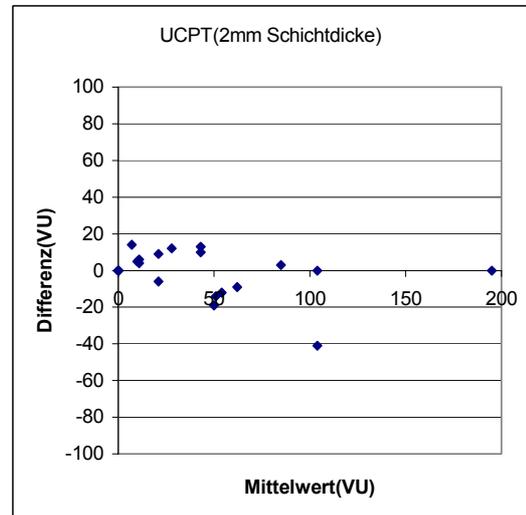
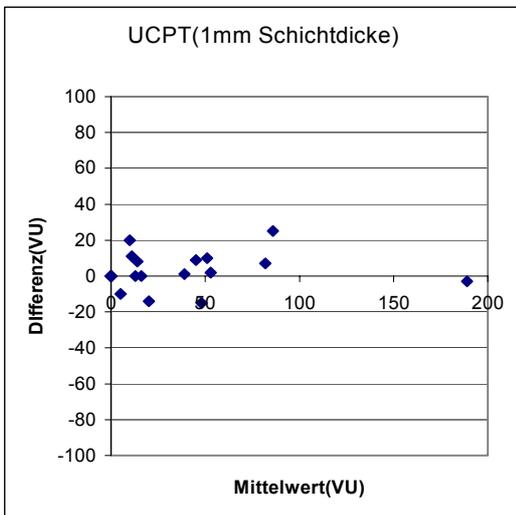
Bei der 2mm-Messung lagen die UCPT-Werte im Bereich zwischen 0VU und 195VU. Die Erstmessung ergab Werte zwischen 0 VU und 195 VU. Bei der Wiederholungsmessung dagegen wurden Werte zwischen 0 VU und 195 VU gemessen. Die mittlere Differenz zwischen erstem und zweitem Wert beträgt 8VU. Die Wertepaarvarianzen variieren zwischen 0VU^2 und 840VU^2 . Der Mittelwert der Wertepaarvarianzen beträgt 71VU^2 .

Wenn man 4mm-Schichten zur Untersuchung der Zirbeldrüsenverkalkung verwendet, liegen die bei unserem Patientenkollektiv gemessenen Werte zwischen 0 VU und 161 VU. Genauer gesagt ergab die erste Messung Meßwerte zwischen 0 VU und 148 VU. Bei der späteren Messung lag der niedrigste Wert bei 0VU und das Maximum bei 161 VU. Die mittlere Differenz zwischen Erstmessung und Wiederholungsmessung beträgt 2,8 VU. Die Wertepaarvarianz schwankt zwischen 0VU^2 und 1568VU^2 . Die mittlere Wertepaarvarianz liegt bei 249VU^2 .

Schließlich fanden wir bei der CT-Messung mit 8mm-Schichten UCPT-Werte von 0VU bis 232VU. Die erste Meßreihe enthielt Werte zwischen 0 VU und 232VU, während die Wiederholungsmessung Werte zwischen 0 VU und 192 VU ergab. Die mittlere Differenz errechnet sich zu 5,7VU. Die Wertepaarvarianz variiert zwischen 0VU^2 und 1265VU^2 , ihr Mittelwert liegt bei 390VU^2 .

Diagrammübersicht 2

Reproduzierbarkeit der UCPT-Messung in Abhängigkeit von der Schichtdicke



In diesen Diagrammen sind die Differenzen der beiden UCPT (=uncalcified pineal tissue)-Messungen gegen den Mittelwert aufgetragen. Hier wird deutlich, daß die Reproduzierbarkeit der UCPT-Messung von der gewählten Schichtdicke abhängig ist. Bei den dickeren Schichten sind die Differenzen größer, d. h., die Reproduzierbarkeit ist schlechter.

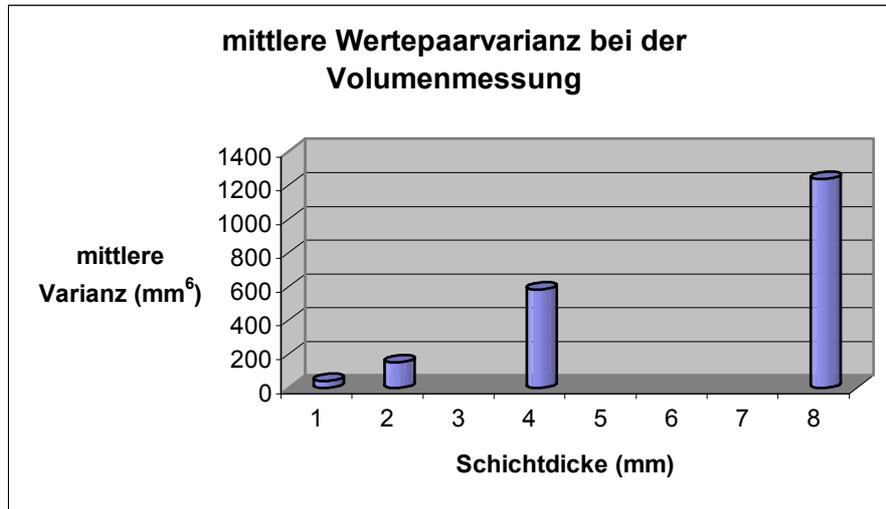


Diagramm 2 Mittlere Wertepaarvarianz der Volumenmessung bei verschiedenen Schichtdicken. In diesem Diagramm wird sichtbar, dass die mittlere Wertepaarvarianz mit zunehmender Schichtdicke ansteigt.

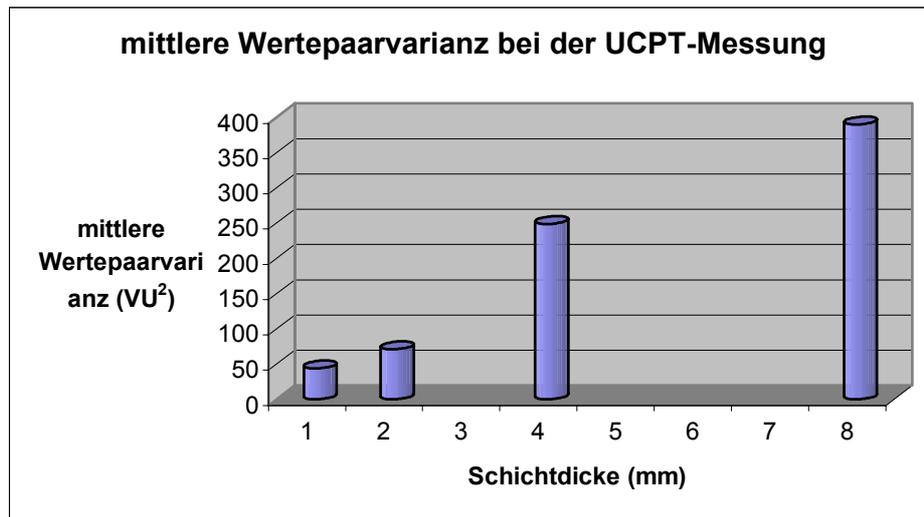
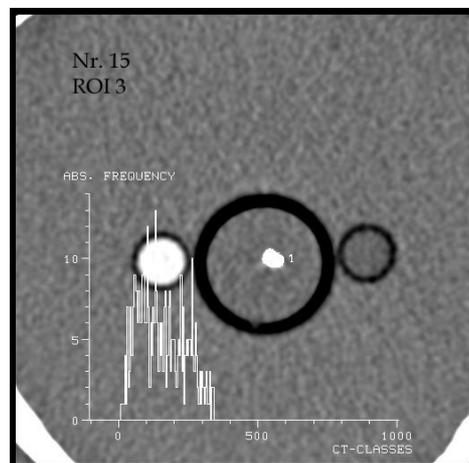
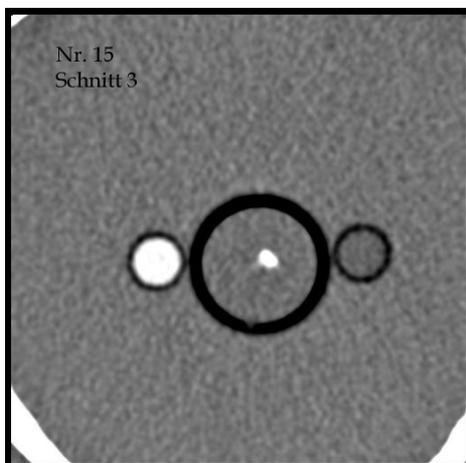
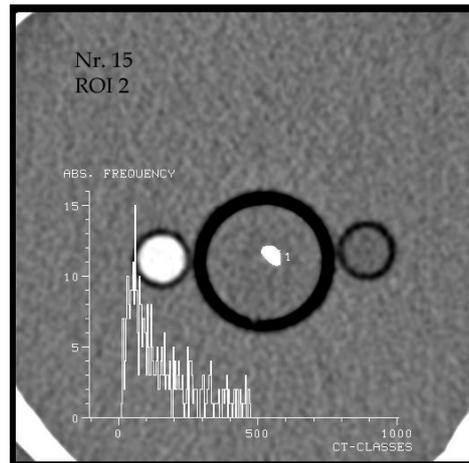
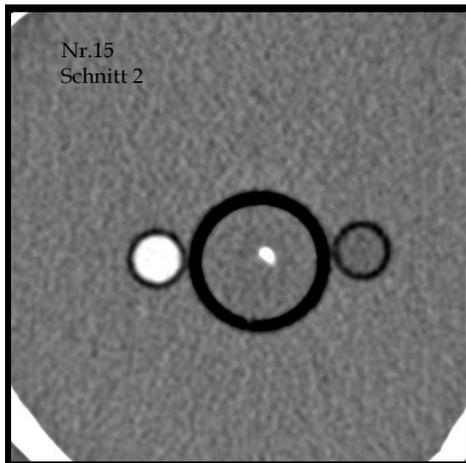
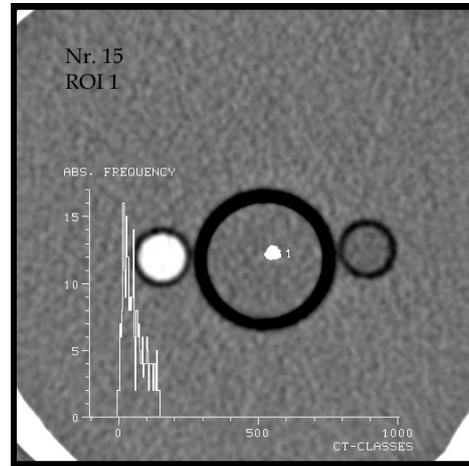
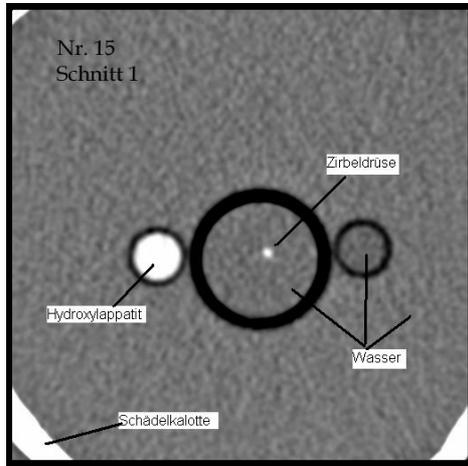
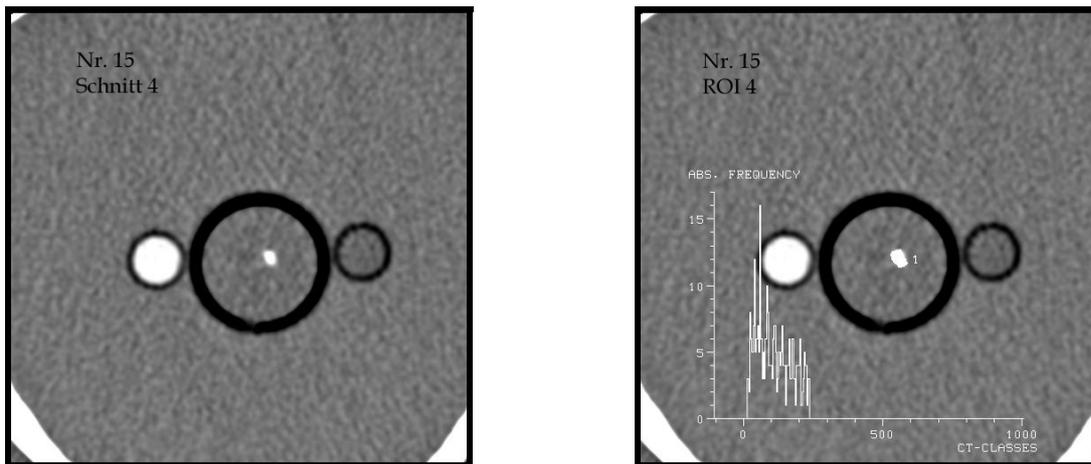


Diagramm 3 Mittlere Wertepaarvarianz der UCPT-Messung bei verschiedenen Schichtdicken. Der bei der Volumenmessung festgestellte Anstieg der mittleren Wertepaarvarianz bei zunehmender Schichtdicke findet sich auch bei der UCPT-Messung.

Bildanhang 1



Anhand dieser Bilder kann das Vorgehen bei der Beurteilung der Zirbeldrüsenverkalkung beispielhaft erläutert werden. Es handelt sich um eine Untersuchung der Zirbeldrüse Nr. 15 mit Hilfe von 1mm-Schichten. Links sind die CT-Aufnahmen des Kopfphantoms, auf denen Teile der Drüse zu sehen sind, rechts sind die dazugehörigen ROIs gezeigt.

Statistische Auswertung der CT-Daten

Die mit Hilfe der Computertomographie gewonnenen Daten können in mehrere Gruppen unterteilt werden. Es wurden zwei verschiedene Größen untersucht: das Drüsenvolumen und das UCPT. Jede dieser Meßgrößen wurde mit Hilfe von vier verschiedenen Meßmethoden untersucht: 1, 2, 4 und 8 mm Schichtdicke. Somit ergeben sich für jede Größe jeweils vier verbundene Stichproben. Die Stichproben sind verbunden, weil immer die gleichen Zirbeldrüsen mit unterschiedlichen Meßmethoden untersucht werden. Diese Tatsache muß bei der statistischen Auswertung unbedingt berücksichtigt werden.

Die statistische Auswertung der oben zusammengefaßten Daten orientiert sich an das von Bland et al. [63] entwickelte Vorgehen. Als primäres Maß für die Reproduzierbarkeit einer Meßmethode dient bei Bland die Standardabweichung der Wertepaar differenzen. Ein daraus abgeleitetes Maß für die Reproduzierbarkeit ist die Wertepaarvarianz. Wenn wie im vorliegenden Fall zwei Messungen pro Objekt gemacht werden, entspricht die Wertepaarvarianz der Hälfte der quadrierten Wertepaar differenz. Während die Standardabweichung der Differenzen besser zur graphischen Darstellung der Reproduzierbarkeit geeignet ist, wird die weniger anschauliche Wertepaarvarianz beim statistischen Vergleich der Reproduzierbarkeit verschiedener Meßmethoden eingesetzt. In diesem Fall wurden die Wertepaarvarianzen der jeweiligen verbundenen Stichproben auf signifikante Unterschiede untersucht. Die Prüfung auf signifikante Unterschiede erfolgte mit Hilfe des Friedman-Tests und des multiplen Vergleichs nach Wilcoxon-Wilcox. Es

handelt sich hier um nichtparametrische Testverfahren. Das genaue Vorgehen ist im Methodenteil und im Anhang 1 detailliert beschrieben.

Zuerst wurden die Volumenmessungen miteinander verglichen:

Der Friedman-Test zeigt, daß signifikante Unterschiede zwischen den Meßmethoden bestehen ($p=0,000$, zweiseitige Fragestellung, $\alpha=0,05$). Daraus können wir aber nicht erfahren, welche Methoden sich voneinander unterscheiden. Der anschließend durchgeführte multiple Vergleich nach Wilcoxon und Wilcox erlaubt den paarweisen Vergleich der einzelnen Methoden.

a) **1mm und 2mm**

$p>0,05$

Kein signifikanter Unterschied ($\alpha=0,05$, zweiseitige Fragestellung).

b) **1mm und 4mm**

$p<0,05$

Signifikanter Unterschied ($\alpha=0,05$, zweiseitige Fragestellung).

c) **1mm und 8mm**

$p<0,05$

Signifikanter Unterschied ($\alpha=0,05$, zweiseitige Fragestellung).

d) **2mm und 4mm**

$p>0,05$

Kein signifikanter Unterschied ($\alpha=0,05$, zweiseitige Fragestellung).

e) **2mm und 8mm**

$p<0,05$

Signifikanter Unterschied ($\alpha=0,05$, zweiseitige Fragestellung).

f) **4mm und 8mm**

$p>0,05$

Kein signifikanter Unterschied ($\alpha=0,05$, zweiseitige Fragestellung).

Nach dem gleichen Muster verfahren wir bei der Auswertung der UCPT-Reproduzierbarkeit. Auch hier werden die Messungen mit Hilfe des Friedman-Tests und des multiplen Vergleichs nach Wilcoxon-Wilcox verglichen. Der Friedman-Test zeigt wieder das Vorhandensein signifikanter Unterschiede zwischen den einzelnen UCPT-Messungen ($p=0,000$, zweiseitige Fragestellung, $\alpha=0,05$). Der multiple Vergleich nach Wilcoxon und Wilcox, der zur genauen Lokalisation der Unterschiede dient, liefert folgende Ergebnisse:

a) **1mm und 2mm**

$p>0,05$

Kein signifikanter Unterschied ($\alpha=0,05$, zweiseitige Fragestellung).

b) 1mm und 4mm $p < 0,05$ Signifikanter Unterschied ($\alpha = 0,05$, zweiseitige Fragestellung).**c) 1mm und 8mm** $p < 0,05$ Signifikanter Unterschied ($\alpha = 0,05$, zweiseitige Fragestellung).**d) 2mm und 4mm** $p > 0,05$ Kein signifikanter Unterschied ($\alpha = 0,05$, zweiseitige Fragestellung).**e) 2mm und 8mm** $p < 0,05$ Signifikanter Unterschied ($\alpha = 0,05$, zweiseitige Fragestellung).**f) 4mm und 8mm** $p > 0,05$ Kein signifikanter Unterschied ($\alpha = 0,05$, zweiseitige Fragestellung).

Hier muß noch einmal unterstrichen werden, daß die hier benutzten nichtparametrischen Tests Ränge und keine bestimmten Parameter, wie z. B. Mittelwerte miteinander vergleichen. Die hier gefundenen signifikanten Unterschiede bedeuten, daß in manchen Stichproben niedrige Werte häufiger auftreten im Vergleich mit anderen Stichproben. Deswegen haben diese Stichproben einen niedrigeren mittleren Rang im Friedman-Test. Wenn der multiple Vergleich nach Wilcoxon-Wilcox signifikante Unterschiede zeigt, so bedeutet das, daß die beiden Stichproben aus unterschiedlichen Grundgesamtheiten stammen und daß eine Stichprobe Meßwerte enthält, die in der Mehrzahl der Fälle kleiner sind als die entsprechenden Meßwerte der anderen Stichprobe. Die oben angegebenen Mittelwerte für jede Meßmethode sollen verdeutlichen, welche Meßmethode (=Stichprobe) die niedrigsten Werte enthält.

3.2. MR-Messung

Ergebnisse

Auf den mit Hilfe der Sequenz ts1 11 erzeugten Bildern waren alle Zirbeldrüsen eindeutig erkennbar. Alle waren hyperintens gegenüber dem umgebenden Wasser, wobei das Ausmaß der Hyperintensität variiert. Pinealiszysten konnten hier nicht gefunden werden. Die Bestimmung des signal intensity ratio ergab SIR-Werte zwischen 0,39 und 0,89, ihr Mittelwert lag bei 0,60. Die Konkremete konnten nicht vom umgebenden Gewebe abgegrenzt werden.

Letzteres gilt auch für die andere T1-gewichtete Sequenz, ts 3d 21. Auch hier waren die Drüsen gut darstellbar. Die SIR-Werte schwanken je nach Zirbeldrüse zwischen 0,36 und 0,83, der Mittelwert liegt bei 0,63.

Bei der T2-gewichteten Sequenz fl 3d 1 war es vor allem bei kleinen Drüsen schwierig, die Grenze zwischen der Drüse und der umgebenden Flüssigkeit zu bestimmen. Innerhalb der Drüsen kann man hellere und dunklere Areale erkennen, die aber offensichtlich in keinem Zusammenhang mit den auf CT-Aufnahmen erkennbaren Konkrementen stehen. Die SIR-Werte liegen hier zwischen 0,48 und 0,87. Der mittlere SIR-Wert beträgt 0,72.

Statistische Auswertung

Bei keiner der hier benutzten Sequenzen konnte ein statistisch signifikanter Zusammenhang zwischen signal intensity ratio (SIR) und dem mittleren Dichtewert der entsprechenden Drüse (gemessen in HU) festgestellt werden.

Bei der Sequenz ts1 11 ergab der Vergleich zwischen den SIR-Werten und den entsprechenden CT-Dichtewerten einen Korrelationskoeffizienten von $r=-0,232$. Der dazugehörige p-Wert ($p=0,299$) zeigt, daß kein signifikanter Zusammenhang zwischen SIR und CT-Dichtewert besteht.

Ähnlich waren die Untersuchungsergebnisse bei der Sequenz ts 3d 21. Zwischen SIR und dem mittleren CT-Dichtewert wurde ein Korrelationskoeffizient von $-0,150$ gefunden. Die anschließende statistische Auswertung zeigt, daß auch hier kein signifikanter Zusammenhang besteht ($p=0,506$).

Auch wenn T2-gewichtete Sequenzen benutzt werden ist kein Zusammenhang zwischen SIR und CT-Dichtewert erkennbar. Bei der Sequenz fl 3d 1 beträgt der Korrelationskoeffizient $r=-0,379$. Der Vergleich mit dem tabellierten kritischen wert ergibt einen p-Wert von 0,082, also besteht auch hier kein statistisch signifikanter Zusammenhang.

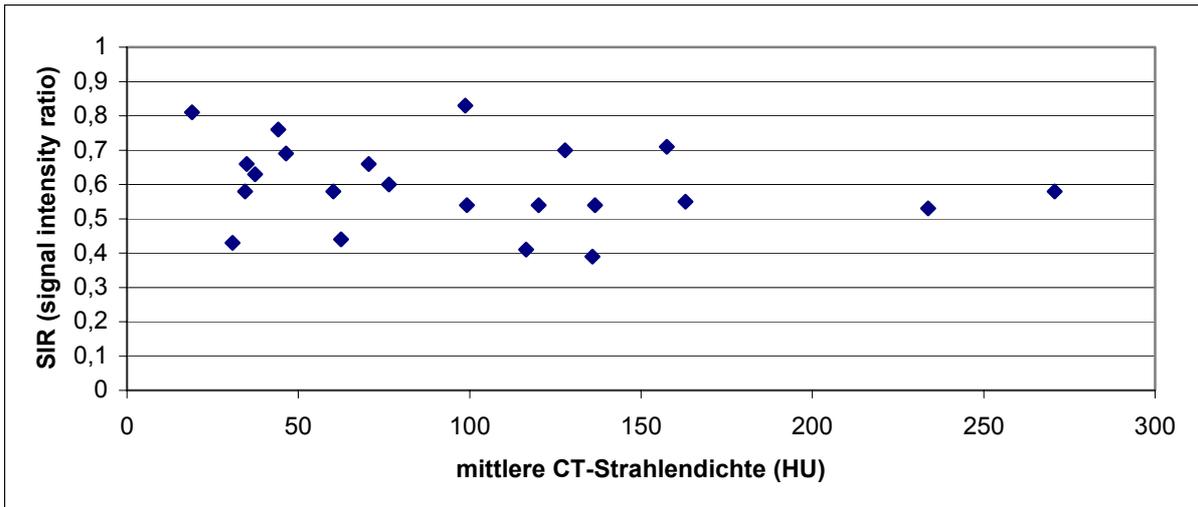


Diagramm 4 In diesem Diagramm ist für alle Drüsen das SIR gegen die mittlere Strahlendichte (in HU) aufgetragen. Das SIR wurde mit Hilfe der Sequenz ts1 11 ermittelt. Es ist keine Abhängigkeit der beiden Variablen zu erkennen.

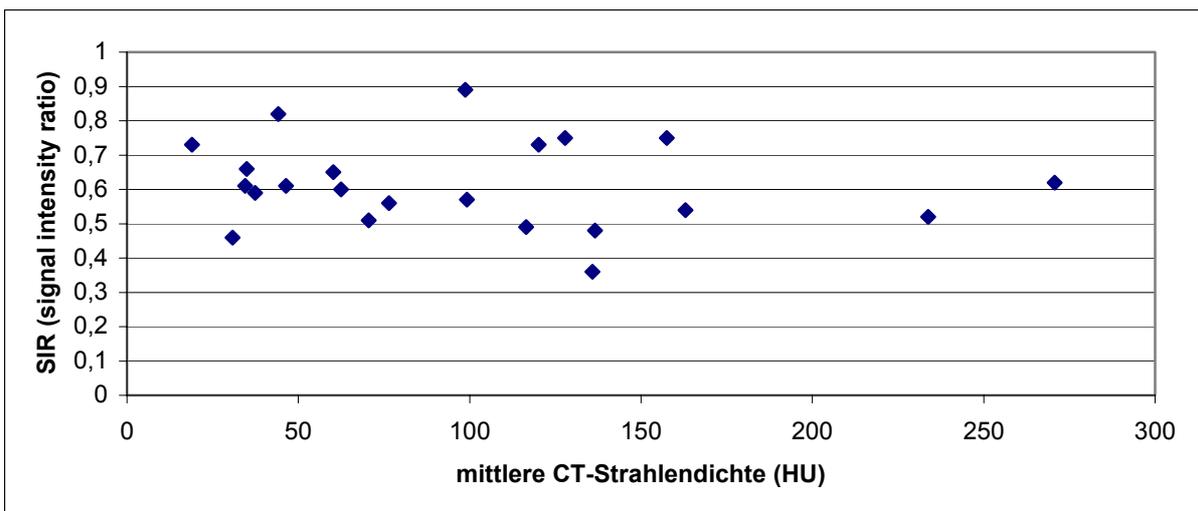


Diagramm 5 In diesem Diagramm ist für alle Drüsen das SIR gegen die mittlere Strahlendichte (in HU) aufgetragen. Das SIR wurde mit Hilfe der Sequenz ts 3d 21 ermittelt. Es ist keine Abhängigkeit der beiden Variablen zu erkennen.

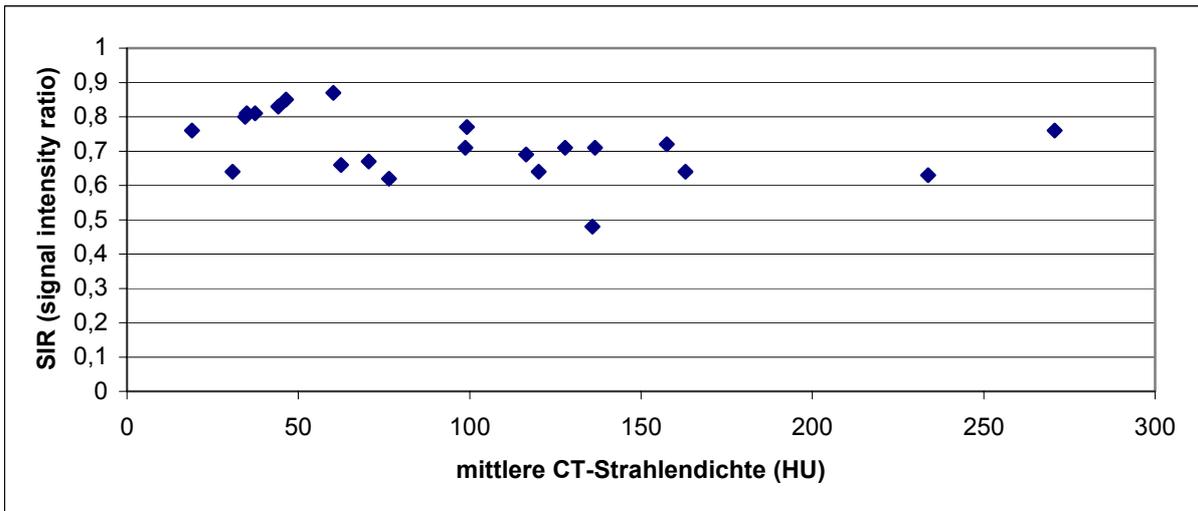
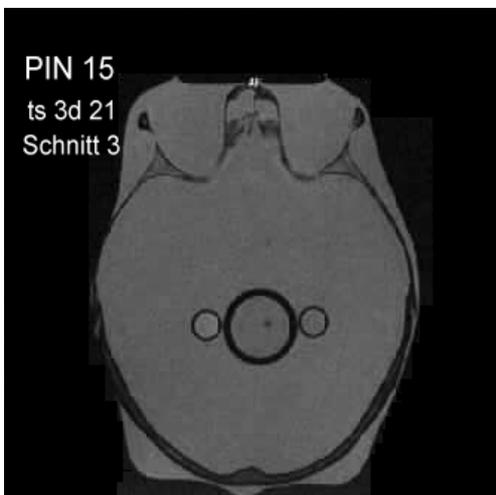
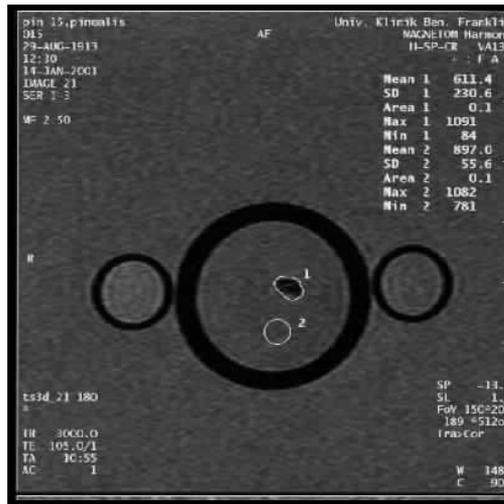
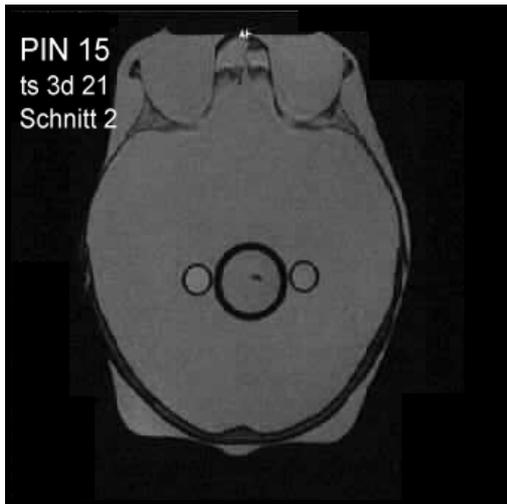
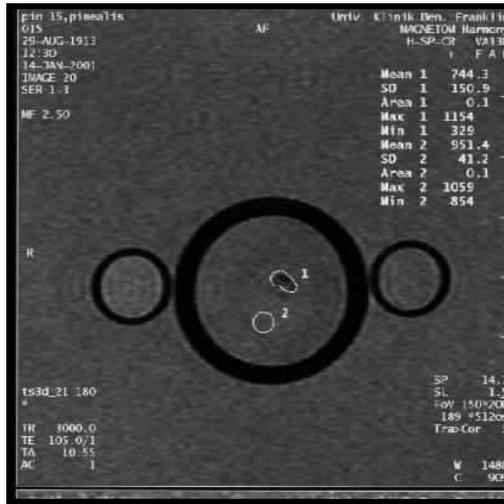
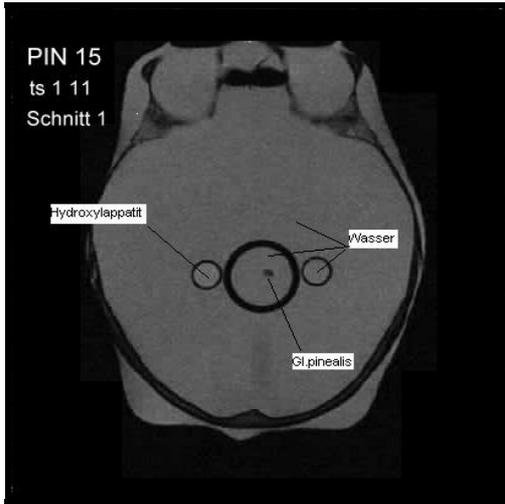
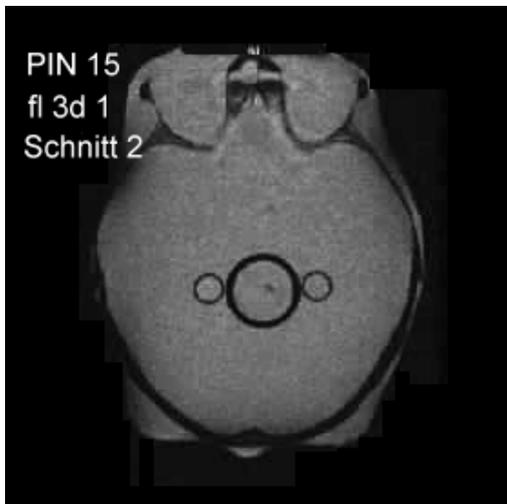
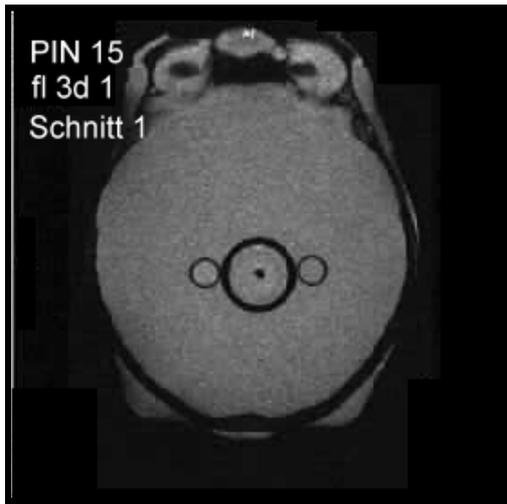
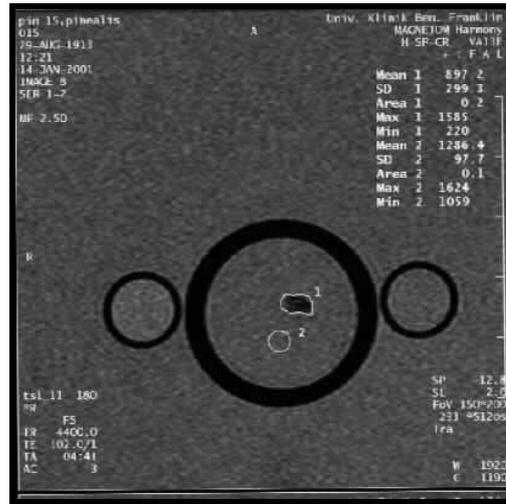
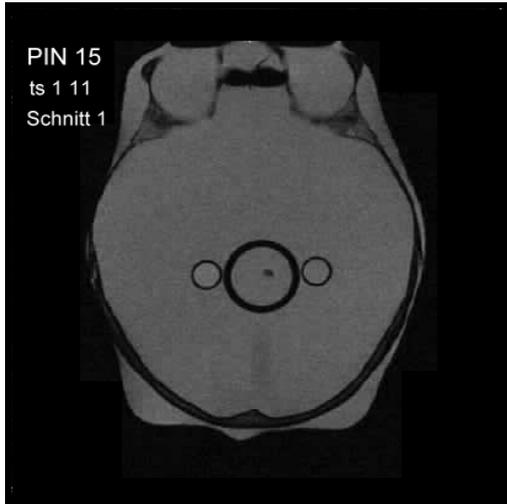


Diagramm 6 In diesem Diagramm ist für alle Drüsen das SIR gegen die mittlere Strahlendichte (in HU) aufgetragen. Das SIR wurde mit Hilfe der Sequenz fl 3d 21 ermittelt. Es ist keine Abhängigkeit der beiden Variablen zu erkennen.

Bildanhang 2a



Bildanhang 2b



In diesem Bildanhang werden MR-Aufnahmen der Drüse Nr.15 mit drei unterschiedlichen Sequenzen gezeigt.