

# Fiabilité et reproductibilité des mesures de volume rénal par IRM

## *Reliability and reproducibility of renal volumetry by MRI*

### CONTEXTE

Le meilleur marqueur de la sévérité de l'atteinte rénale dans la polykystose est l'augmentation du volume des reins. L'IRM est la méthode d'imagerie de choix pour mesurer ces volumes, le CHUV l'utilise depuis des années pour le suivi clinique de ses patients.

### OBJECTIFS

La fiabilité de la volumétrie rénale est essentielle pour assurer une prise en charge optimale des patients. Nous avons comparé la précision et la reproductibilité des diverses techniques de volumétrie par IRM et évalué la reproductibilité intra- et inter-observateur de la technique utilisée en clinique.

### MÉTHODE

Nous avons comparé les différentes techniques disponibles en utilisant un fantôme et en lui appliquant le protocole d'imagerie IRM *rein polykystique* utilisé pour les patients. Nous avons étudié la reproductibilité intra- et inter-observateur en reprenant les dossiers de 37 patients et en mesurant leurs volumes rénaux.

### RÉSULTATS

La technique la plus fiable et reproductible déterminée sur le fantôme était celle de seuillage manuel d'AW Server sur des coupes transverses en pondération T2. Sur les patients, la variabilité intra-observateur était  $< 2\%$ , et l'inter-observateur de  $< 4\%$ . Les limites d'agrément étaient de  $-70 ; 90$  ml tant pour la variabilité intra- qu'inter-observateur.

### CONCLUSION

Les différentes méthodes évaluées sont reproductibles, mais leur précision est très variable. Il y a des biais de mesure systématiques avec toutes les techniques. Nos mesures montrent une haute reproductibilité intra-observateur et inter-observateurs. Nous avons identifié une limite d'agrément supérieure de 90 ml, ce qui signifie que des variations de volume au cours du temps de  $< 90$  ml ne peuvent pas être considérées comme significatives. Des augmentations de  $> 90$  ml ou des régressions de  $> 70$  ml peuvent être interprétées comme réelles de façon fiable avec la volumétrie IRM.

**MOTS-CLÉS** : volumétrie, IRM, polykystose, ADPKD

Mémoire de Maîtrise en médecine n°883

# **Fiabilité et reproductibilité des mesures de volume rénal par IRM**

## **Étudiant**

Guillaume Wegmüller

## **Tuteur**

Prof Jean-Yves Meuwly  
Service de Radiodiagnostic, CHUV

## **Expert**

PD Dr med Andreas Serra  
Dpt de néphrologie, Zürich

Lausanne, janvier 2015

# TABLE DES MATIÈRES

CONTEXTE.....	1
OBJECTIFS .....	2
MATÉRIEL ET MÉTHODE.....	3
Le fantôme – construction.....	3
Le fantôme – détermination du volume par déplacement des fluides.....	3
Le fantôme – IRM.....	3
Les patients .....	4
Les patients – variabilité intra-observateur .....	4
Les patients – variabilité inter-observateurs .....	5
Analyse statistique .....	5
RÉSULTATS .....	6
Construction du fantôme.....	6
Immersion du fantôme .....	6
Volumétrie du fantôme.....	6
Mesures sur les patients (intra-observateur) .....	7
Mesures sur les patients (inter-observateurs).....	7
DISCUSSION .....	8
BIBLIOGRAPHIE.....	13
ANNEXES .....	15

## CONTEXTE

La polykystose rénale autosomique dominante (ou ADPKD, de l'anglais *autosomal dominant polycystic kidney disease*) est la plus courante des maladies héréditaires monogéniques du rein, son incidence est estimée à environ une naissance sur 800 (1). Deux gènes responsables ont été identifiés, PKD-1 et PKD-2, codant pour les protéines de polycystine. Ces protéines sont essentielles pour maintenir la polarisation de certaines cellules du corps. Les atteintes des gènes PKD-1 ou PKD-2 conduisent à la transformation de structures normalement tubulaires en structures kystiques. La maladie qui en découle se caractérise avant tout par la croissance de ces kystes à partir des tubules rénaux, mais on retrouve des kystes dans d'autres organes, en particulier le foie, les vésicules séminales, voire le pancréas et les plexus arachnoïdes (2, 3). La polykystose rénale possède aussi des manifestations non-kystiques (hypertension artérielle, anévrysmes, valvulopathies cardiaques, bronchiectasies, hernies inguinales bilatérales). La morbidité et la mortalité de l'ADPKD sont liées aux complications rénales et cardiovasculaires. Bien que de nombreuses recherches soient menées, il n'existe à l'heure actuelle aucun traitement curatif, les options thérapeutiques reposent sur la limitation des complications et sur le ralentissement de la croissance des kystes.

Au niveau rénal, les kystes compriment et détruisent le parenchyme normal par leur croissance, menant à un élargissement de cet organe et contribuant à son dysfonctionnement. Dans plus de la moitié des cas, la maladie progresse jusqu'à l'insuffisance rénale terminale, dont elle est la quatrième cause la plus fréquente en Suisse avec 11% des patients en dialyse souffrant d'ADPKD (4-6). Il est difficile de déterminer quand un patient arrive en insuffisance rénale dans le cadre de la polykystose : l'élévation de la créatininémie, qui est habituellement utilisée en clinique pour estimer le débit de filtration glomérulaire et évaluer la fonction rénale, n'apparaît que très tardivement dans le cours de la maladie. En effet, les kystes se développent dans une minorité des tubules rénaux, et les glomérules sains sont longtemps capables de compenser la dysfonction des glomérules atteints. Ainsi l'insuffisance rénale ne survient en moyenne qu'à partir de 30 ou 40 ans d'évolution (7). C'est pourquoi les efforts des chercheurs se sont concentrés ces dernières années sur la recherche d'autres marqueurs, plus précoces, permettant de surveiller l'évolution de l'atteinte rénale et de suivre l'effet des thérapies.

Un groupe de chercheurs, le CRISP (pour *Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease*), a été créé en 2000 afin de développer des méthodes radiologiques pour évaluer la progression de la maladie. Ces chercheurs ont démontré que la sévérité de l'atteinte rénale dans l'ADPKD est directement associée à l'augmentation du volume des reins (8). En parallèle, le CRISP et d'autres groupes de chercheurs ont comparé les performances de diverses modalités d'imagerie pour la mesure des volumes rénaux. L'échographie, qui joue un rôle majeur pour poser le diagnostic de polykystose rénale (9, 10), est peu fiable quand il s'agit de mesurer des volumes. Le CT-scan et l'imagerie par résonance magnétique (IRM) sont tous deux efficaces, mais l'IRM présente de nombreux avantages qui lui ont permis de rapidement s'imposer comme la technique radiologique de prédilection (11). En effet, elle offre une meilleure résolution des tissus mous, elle est non-irradiante, et ne nécessite pas l'emploi de produits de contraste iodés néphrotoxiques ni de gadolinium (12, 13).

## **OBJECTIFS**

La volumétrie par IRM est utilisée au Centre Hospitalier Universitaire Vaudois (CHUV) pour le suivi des patients atteints de polykystose hépatorénale. La fiabilité de ces mesures est essentielle pour assurer une prise en charge optimale de ces patients. L'objectif de cette étude était de comparer la précision et la reproductibilité des diverses techniques de volumétrie par IRM en utilisant un fantôme, d'évaluer la reproductibilité intra-observateur de la technique utilisée en clinique et de vérifier si les mesures effectuées sur des reins polykystiques depuis l'introduction de ces protocoles IRM en pratique clinique sont reproductibles.

## **MATÉRIEL ET MÉTHODE**

Cette étude est divisée en deux parties : la première partie de l'étude a été pratiquée sur un fantôme simulant un rein polykystique, le but étant de déterminer quelle méthode de volumétrie parmi celles disponibles dans le service de radiodiagnostic du CHUV était la plus fiable et la plus reproductible. Dans la seconde partie, les volumes rénaux droits et gauches de patients suivis au CHUV ont été mesurés avec la méthode utilisée en clinique. Ces mesures ont été comparées avec les volumes obtenus à l'origine par les collaborateurs du service d'imagerie avec cette méthode. Ces comparaisons ont été utilisées afin de déterminer la variabilité inter-observateur. Pour évaluer la variabilité intra-observateur, un examinateur a mesuré trois fois à quelques jours d'intervalle le volume des reins à partir d'une même série d'images, et ce chez cinq patients.

Cette étude a été approuvée par le Comité d'éthique institutionnel. L'analyse des données des patients a été réalisée de manière rétrospective. Aucun consentement n'a été requis.

### ***Le fantôme – construction***

Un fantôme simulant un rein polykystique a été construit avec six sphères (4 sphères de 5 centimètres de diamètre, une sphère de 6 centimètres et une de 7 centimètres) et deux ellipsoïdes (7 centimètres de diamètre et 3 centimètres d'épaisseur) en plastique, remplis d'eau, qui représentaient les kystes rénaux (Figure 1). Chaque élément était formé de deux pièces s'emboîtant. Le volume calculé des sphères était de 65.42, 113.04 et 179.15 millilitres respectivement. Le volume d'un ellipsoïde était de 76.78 millilitres. Le volume calculé du fantôme était de 707.43 millilitres. La paroi des éléments mesurait 0.2 millimètres. Ces huit "kystes" ont été pesés à cinq reprises sur une balance Mettler-Toledo PB1502-S (précision 0.01 gramme). Ces éléments ont été collés ensemble afin de les assembler avec une morphologie de rein polykystique.

### ***Le fantôme – détermination du volume par déplacement des fluides***

Pour mesurer le volume du fantôme, un bac rempli d'eau déminéralisée et percé à mi-hauteur par une sonde d'aspiration a été utilisé, de sorte que toute élévation du niveau de l'eau dans le bac conduisait à l'aspiration de l'excédent de liquide en direction d'un récipient gradué (Figures 2). Ce récipient était posé sur une balance Mettler-Toledo PB1502-S tarée, de façon à pouvoir mesurer le poids du liquide déplacé, en sus de son volume.

Le principe d'Archimède veut que le volume du corps immergé soit égal au volume de fluide déplacé. La quantité d'eau aspirée vers le récipient gradué correspond donc au volume des objets plongés dans le bac. Le bon fonctionnement du montage a été validé en réalisant quelques essais d'immersion de briques de boisson d'un volume connu (Tetrabricks de 250 et 500 ml). Le vacuum d'aspiration a été arbitrairement fixé à -30 kPa. Quatre mesures du volume du fantôme ont été réalisées. Afin d'obtenir un état d'équilibre, les mesures ont été lues 30 minutes après que le fantôme a été plongé dans le liquide. Le volume a été déterminé sur la base des graduations du récipient.

### ***Le fantôme – IRM***

Le modèle de rein polykystique a ensuite été immergé dans un bac de 4 litres rempli de saindoux liquide qui s'est solidifié en refroidissant. Le saindoux a permis de mimer la graisse péri-rénale

(Figure 3).

Le fantôme a ensuite été passé dans deux machines d'IRM (Siemens Verio 3T et Siemens Aera 1.5T) (Figure 4). Les protocoles *rein polykystique* utilisés pour le suivi des patients ont été employés sur les deux machines. Il s'agissait d'une série T2 transverse HASTE (écho de spin ultra rapide) avec des coupes de 4 mm d'épaisseur, d'une série T2 HASTE coronale à 6 mm d'épaisseur, et d'une série T1 coronale à 6 mm d'épaisseur (Figure 5).

Le volume du fantôme sur les six séries d'images a été déterminé à l'aide de deux programmes informatiques, Advantage Workstation Server de GE Healthcare (AW Server) et Vue PACS de CareStream (CS). Chacune des six séries a été analysée par un seul observateur à vingt reprises avec cinq techniques de mesures : main levée, contour automatique et seuillage d'AW Server, main levée et remplissage automatique de CareStream (Figure 6).

Les procédures à main levée consistaient à tracer le contour du fantôme sur les coupes successives, le logiciel calculant ensuite le volume total en additionnant les aires délimitées sur chaque coupe et les multipliant par l'épaisseur de la coupe.

Les trois techniques semi-automatiques se basaient sur le fait que les différents tissus sont caractérisés par une intensité différente et qu'on peut donc séparer les divers pixels formant une image selon leur intensité. Le seuillage permettait de sélectionner certains tissus aux dépens d'autres en appliquant un seuil d'intensité de gris, le programme ne conservant que les pixels ayant une intensité supérieure au seuil. Dans le cas du fantôme, on pouvait ainsi séparer la graisse des structures aqueuses mimant les kystes en appliquant un seuil qui ne sélectionnait que ces dernières. Le contour automatique fonctionnait d'une manière similaire, en détectant les limites entre des tissus et structures d'intensités différentes. Pour le remplissage automatique, il fallait déterminer sur chaque coupe un ou plusieurs points d'intérêt. On étendait ensuite la sélection aux pixels adjacents qui avaient la même intensité que le point d'origine. Dans les trois cas, le programme multipliait ensuite la surface ainsi choisie par l'épaisseur de la coupe et sommait les volumes de chaque coupe pour donner le volume total de l'objet d'intérêt.

## **Les patients**

Après avoir comparé les méthodes volumétriques à l'aide du fantôme, la seconde étape de l'étude a été réalisée sur les séries d'images de patients. Les dossiers de 37 patients consécutifs souffrant de polykystose rénale et qui avaient bénéficié d'une IRM au CHUV entre juillet 2008 et novembre 2011 pour l'évaluation du volume des reins ont été utilisés (19 hommes, 18 femmes, âge moyen 39 ans, écart 18-69 ans). Le volume des reins de ces patients avait été calculé par des opérateurs différents avec la technique de contour à main levée du logiciel AW sur les coupes transverses ou coronales T2 de machines 1.5 Tesla ou 3 Tesla (27 à 1.5T, 10 à 3T). Pour effectuer la nouvelle mesure des volumes rénaux de ces patients, la technique de volumétrie par contour à main levée du logiciel AW Server a été employée. Les images des séquences T2 transverses ont été utilisées (Figure 7).

## **Les patients – variabilité intra-observateur**

Parmi les 37 patients, 5 patients ont été choisis aléatoirement pour la mesure de la variabilité intra-observateur. Un examinateur entraîné (GW) a mesuré trois fois entre 2 et 9 jours d'intervalle le

volume des deux reins de ces 5 patients.

### ***Les patients – variabilité inter-observateurs***

Un examinateur (GW) a mesuré une fois le volume rénal gauche et droit de chacun des 37 patients. Les résultats des nouvelles mesures ont été comparés aux volumes de référence mesurés par les médecins du service de radiodiagnostic du CHUV.

Les contours du rein droit étant souvent plus difficiles à discerner que ceux du rein gauche en cas de polykystose, à cause de l'imposante présence du foie et de ses nombreux kystes, la reproductibilité de la volumétrie à droite par rapport à gauche a été évaluée.

### ***Analyse statistique***

La variabilité des mesures de volume du fantôme avec les différentes machines, les différentes séquences et les différents programmes a été analysée selon la technique de Bland-Altman (14). Les box-plots permettant de comparer les mesures ont été effectués avec le programme statistique Stata 13.1 (StataCorp, College Station). La fiabilité des mesures a été définie comme le décalage en pourcent par rapport à la mesure de référence réalisée avec la technique de déplacement des volumes. La concordance entre les mesures effectuées sur les deux machines a été analysée en termes de pourcentage. L'étude de la variabilité intra- et inter-observateur pour les mesures cliniques a été effectuée avec l'analyse de Bland-Altman.

Pour les mesures cliniques, la variabilité inter-observateurs a également été effectuée à l'aide de plusieurs indices de concordance : le coefficient de variation, le coefficient de corrélation (precision) de Pearson, et le coefficient d'exactitude (accuracy) de Lin. Les valeurs ont été calculées avec le programme statistique Stata 13.1 (StataCorp, College Station). Les indices de concordance des mesures réalisées sur le rein droit ont été comparés avec les indices des volumes du rein gauche. Les indices de concordance des mesures effectuées sur des reins de petit volume ont également été comparés avec ceux obtenus sur des reins de gros volume. La limite entre les deux sous-groupes a été fixée à 600 ml.

# RÉSULTATS

## ***Construction du fantôme***

Les résultats des mesures de la masse de notre fantôme sont représentés dans le Tableau 1. Les 8 kystes pesaient 715.1 g. Après assemblage, le fantôme pesait 717.4 g, les 2.3 grammes de différence étant dus à la colle à prise rapide.

## ***Immersion du fantôme***

Le Tableau 2 montre les résultats des mesures par immersion du fantôme. La première mesure avait été obtenue après un temps d'attente de 15 minutes. La moyenne des trois mesures suivantes a été prise en compte pour déterminer le volume réel du fantôme, soit 715 ml. Cette valeur est utilisée comme référence pour la suite des mesures du fantôme.

## ***Volumétrie du fantôme***

Les résultats complets de volumétrie sur le fantôme sont illustrés dans les Figures 8 et 9. Sur les séquences pondérées en T1, il n'a pas été possible d'obtenir des mesures avec les options mesure automatique et seuillage du logiciel AW Server.

Les résultats des mesures de volumes sur Aera 1.5T étaient très variables en fonction des séquences utilisées et des méthodes d'analyse (Tableau 3). La technique la plus reproductible (T2 trans AW auto) sous-estimait considérablement le volume (biais de -60.95 ml pour une déviation standard de 2.80 ml). La technique qui donnait le meilleur résultat moyen (T2 cor CS auto) avait par contre une grande dispersion des résultats (biais de -3.35 ml pour une déviation standard de 52.13 ml).

Les résultats des mesures de volumes sur Verio 3T étaient également très variables en fonction des séquences utilisées et des méthodes d'analyse (Tableau 3). La technique la plus reproductible (T2 cor AW auto) sous-estimait le volume (biais de -42.65 ml pour une déviation standard de 2.78 ml). La technique qui donnait le meilleur résultat moyen (T2 cor AW main) avait par contre une importante dispersion des résultats (biais de 13.82 ml pour une déviation standard de 11.22 ml).

La Figure 10 représente la fiabilité des diverses méthodes. Les méthodes semi-automatiques étaient systématiquement plus fiables en coupes coronales qu'en coupes transverses. Le phénomène inverse était observé avec les méthodes manuelles : les volumes mesurés sur les coupes transverses étaient plus proches de la référence que les volumes mesurés sur les coupes coronales.

En comparant les volumes mesurés sur les images fournies par l'IRM Aera 1.5T avec ceux mesurés sur les images de l'IRM Verio 3T, on remarque que les écarts de concordance étaient entre 0.1 % (CS main T2 transverse) et 5.9% (CS auto T2 transverse) (Tableau 4). La variation était de 2.5% en moyenne pour les techniques manuelles contre 3.5% pour les techniques semi-automatiques.

L'analyse de Bland-Altman montrait que l'erreur systématique (biais) était plus importante avec la machine Verio que Aera (moyenne 0.86 pour Aera, -14.51 pour Verio, T-Test  $p < 0.05$ ). La technique de mesure CS automatique des coupes coronales T2 sur la machine Aera donnait un résultat précis (biais = -3.35) mais avec la plus grande dispersion (limits of agreement -107.61, 100.91). La technique de mesure la plus fiable était la mesure AW avec seuillage sur les coupes transverses T2 sur la machine Aera. Cette technique sous-estimait légèrement la mesure de volume (biais = -8.15)

avec une faible dispersion (limits of agreement -25.19, 8.88). Ceci signifiait que la différence entre deux mesures faites avec cette technique devait être inférieure à -25.19 millilitres ou supérieure à 8.88 millilitres pour être considérée comme significative.

### ***Mesures sur les patients (intra-observateur)***

La volumétrie rénale avec la technique de seuillage manuel AW sur les coupes transverses était reproductible (Tableau 5). La variation entre plusieurs mesures d'un même volume à plusieurs jours d'intervalle était inférieure à 2%. Le coefficient de corrélation de Pearson était compris entre 0.9960 et 0.9998.

L'analyse de Bland-Altman de ces mesures est illustrée sur la Figure 11.

La variabilité intra-observateur montrait une discrète erreur systématique de surestimation (7.79). Les limites d'agrément (limits of agreement = LoA) inférieures à 90cc indiquaient qu'une différence de mesure supérieure à 90cc lors de mesures successives devait être considérée comme significative. Une différence de moins de 90cc devait être considérée comme liée à la variabilité intra-observateur. En séparant l'analyse des reins en fonction de leur volume, en plaçant un seuil à 600 millilitres, le biais de mesure augmentait avec les reins de petit volume (11.85 ml) et diminuait pour les reins de volume supérieur à 600 millilitres (5.09 ml). La LoA passait à -18.25;42.05 pour les reins <600ml et -94.79;104.97 pour les reins >600ml. La représentation graphique de l'analyse de Bland-Altman en fonction du volume des reins apparaît dans la Figure 12.

### ***Mesures sur les patients (inter-observateurs)***

Le Tableau 6 présente les indices de concordance entre les volumes mesurés en clinique et ceux de notre étude. La concordance moyenne entre les volumes de référence et nos mesures était de 95.2%, avec un coefficient de variation de 3.9%. Les coefficients de Pearson et de Lin montraient une concordance proche de 1 (0.997). Il y avait peu de variabilité entre les mesures effectuées pour l'étude et celles réalisées par les médecins en routine clinique (Figure 13).

L'analyse de Bland-Altman des mesures effectuées sur les patients est illustrée sur la figure 14.

La variabilité inter-observateur montrait une discrète erreur systématique de surestimation (8.84). Le coefficient de répétition (limits of agreement) égal à 90cc indique qu'une différence de mesure supérieure à 90cc lors de mesures successives doit être considérée comme significative. Une différence de moins de 90cc doit être considérée comme liée à la variabilité inter-observateur. C'est le même seuil de signification que celui qui avait été identifié pour la variabilité intra-observateur.

Le Tableau 7 indique la reproductibilité des mesures sur le rein droit par rapport au rein gauche, ainsi que la reproductibilité des mesures entre les reins relativement préservés structurellement et les reins déjà très élargis et déformés. Ces tableaux indiquent les évaluations réalisées avec le calcul des coefficients de corrélation de Pearson et de Lin.

Les indices de concordance étaient identiques entre droite et gauche. La concordance était un peu moins bonne en cas de volumétrie sur des petits reins par rapport aux reins massivement agrandis par la maladie.

En étudiant les mêmes données avec l'analyse de Bland-Altman, des différences importantes apparaissaient entre les deux groupes. La Figure 15 présente cette analyse sous forme graphique.

## DISCUSSION

La polykystose hépatorenale est une maladie fréquente qui touche environ 10'000 patients en Suisse (15). Cette maladie a des répercussions importantes, la principale étant l'insuffisance rénale conduisant à la dialyse ou nécessitant une transplantation. Cette insuffisance rénale a tendance à rester longtemps masquée dans le cours de la maladie, elle devient manifeste en général après plusieurs décennies d'évolution, mais l'âge d'apparition est extrêmement variable et imprévisible. Les marqueurs habituels de dégradation de la fonction rénale ne sont pas fiables en ce qui concerne l'ADPKD. Le volume rénal total et son augmentation au cours du temps se sont révélés être de bons marqueurs de substitution, et l'IRM la meilleure modalité d'imagerie pour les mesurer. Plusieurs études ont montré la fiabilité de cette technique pour mesurer le volume des reins polykystiques et suivre de manière précise leur évolution.

Notre étude montre que les différentes techniques de mesure testées étaient reproductibles, mais que leur précision était très variable. Des biais de mesure apparaissaient systématiquement avec toutes les techniques, mais étaient moins importants avec la machine à 1.5T qu'avec la machine à 3T. Certaines techniques de mesure produisaient des biais de l'ordre de 10%, soit en surestimation, soit en sous-estimation du volume. Les limites d'agrément étaient également très variables, avec des techniques dispersant les valeurs de mesures dans des écarts jusqu'à 25% de la valeur du volume mesuré.

Lors de l'évaluation de la variabilité intra et inter-observateur à l'aide de la technique la plus fiable identifiée préalablement (méthode de délimitation à main levée des coupes transverses T2 sur le logiciel AW Server), un biais constant de 8 millilitres environ était observé sur les mesures, avec une limite d'agrément de 90 millilitres dans le sens positif. Ces valeurs étaient modifiées en analysant séparément les reins selon un seuil de 600 millilitres. La limite d'agrément descendait à environ 40 millilitres pour les reins <600 millilitres et montait à 110 millilitres pour les reins >600 millilitres.

La méthode semi-automatique de seuillage d'AW Server était la plus reproductible pour mesurer le volume du fantôme. Avec cette méthode le programme mesurait le volume automatiquement à partir d'un seuil d'intensité décidé au départ par l'œil de l'utilisateur. Le processus était répété sur chaque coupe et le programme faisait enfin la somme du volume choisi sur chaque coupe, en fonction de l'épaisseur de coupe. Dans la situation du fantôme, où les limites d'intensité entre l'objet à segmenter et la graisse étaient très nettes, il n'y avait pas d'ambiguïté dans le choix du seuil. Cette technique sous-évaluait toutefois systématiquement le volume du fantôme, peut être en raison d'une trop grande sensibilité aux pixels d'intensité différente du liquide et de la graisse, comme les quelques bulles d'air résiduelles dans les sphères constituant le fantôme ou les parois des sphères. En situation clinique, où les différences d'intensité entre les organes sont plus variables, le choix du seuil pourrait être plus aléatoire, rendant cette technique difficile à utiliser. La grande variabilité des mesures avec la méthode de remplissage automatique de CareStream, dont les performances sont nettement en-dessous des autres, pourrait s'expliquer par la complexité d'utilisation de la technique.

Sur le plan de la fiabilité, on se rend compte que les trois méthodes semi-automatiques étaient plus reproductibles en coupes coronales qu'en coupes transverses. Cela provient probablement du fait que l'entier du fantôme pouvait être exploré sur moins de 20 coupes coronales, contre plus de 50 en transverse. L'augmentation du nombre de coupes augmentait le risque d'imprécisions et d'erreurs

d'interprétation de la part du logiciel, notamment au niveau des artefacts (présence de quelques bulles dans la graisse, visible seulement sur les coupes transverses de 4 millimètres d'épaisseur et que les algorithmes automatisés n'étaient pas capables de différencier des kystes du fantôme). La situation était inverse pour les mesures faites manuellement : les volumes mesurés sur les coupes transverses étaient les plus proches de la réalité. La délimitation précise des contours du fantôme sur chaque coupe excluait volontairement les bulles d'air situées en dehors du fantôme. Une erreur de délimitation de quelques pixels sur les coupes transverses de 4 millimètres avait moins de répercussion que sur les coupes coronales de 6 millimètres. Le même écart de quelques millimètres en dessinant le contour d'un rein représentait un volume 50% plus important en coronal. L'effet de volume partiel était également moins important sur les coupes d'entrée et de sortie du volume du rein sur les coupes transverses que sur les coupes coronales, car la surface de la coupe était plus faible sur les coupes transverses que sur les coupes coronales.

Le volume théorique du fantôme était de 707.43 millilitres, alors que le volume déterminé par la méthode de déplacement des fluides était de 715 millilitres. Cette différence pouvait être due à l'imprécision du moulage des éléments plastiques, à la colle (2.3 grammes) et à l'épaississement de la coque à la jonction entre deux éléments plastiques constituant une sphère ou un ellipsoïde. Comme la volumétrie IRM était réalisée en mesurant le contenu des éléments de plastique, une sous-estimation du volume devait être attendue. Avec des parois de 0.2 millimètres, elle est devant être de l'ordre de 1%, soit un volume mesuré en volumétrie IRM de 707 millilitres. Compte tenu de cette correction, la technique la plus fiable reste celle identifiée préalablement (méthode de délimitation à main levée des coupes transverses T2 sur le logiciel AW Server). Le biais négatif constant de 8 millilitres environ qui était observé sur les mesures correspondait à la réalité du volume d'eau contenu dans le fantôme.

Les mesures obtenues sur la machine à 1.5 Tesla étaient plus précises que celles obtenues sur la machine à 3 Tesla. Avec un champ plus intense et homogène, la machine à 3 Tesla aurait dû fournir au contraire de meilleurs résultats. Le fantôme était immergé dans un récipient de 4 litres contenant de la graisse. Celui-ci ne représentait pas le volume d'un corps humain et il est possible que les inhomogénéités de champ engendrées par le fantôme lui-même aient provoqué des distorsions responsables de la variabilité des mesures. Par ailleurs, un effet de "chemical-shift" apparaissait dans la direction de la fréquence avec les deux champs utilisés, mais était plus marqué à 3 Tesla qu'à 1.5 Tesla. La mauvaise définition des contours engendrée par cet artefact a pu avoir des répercussions importantes sur la précision des mesures, particulièrement à 3 Tesla.

Bae et al. (16) avaient comparé les mesures obtenues sur un fantôme avec des séquences T2 transverses obtenues avec deux différentes machines à 1.5 Tesla. Ils avaient observé une très bonne reproductibilité avec des variations de l'ordre de 2%. Ils n'avaient toutefois réalisé que deux mesures sur chaque image et leur analyse statistique se limitait à la mesure de la variation en pourcent par rapport au volume connu du fantôme déterminé par la méthode de déplacement des fluides. Notre étude a évalué cinq méthodes différentes de mesures pour des images axiales T2, coronales T1 et T2, réalisées à 1.5 et 3 Tesla. Chaque série a été mesurée 20 fois. Nous n'avons pas retrouvé d'évaluation aussi extensive dans la littérature.

Mis à part les études de Bae et al. (17, 18), la méthodologie utilisée pour évaluer la fiabilité a toujours été la corrélation de Pearson ou de Lin, que ce soit pour les études de fantômes ou les études sur patients. Ces méthodes de corrélation indiquent le degré de relation linéaire entre deux

variables. Dans les études de volumétrie rénale, c'est toujours le même volume qui est mesuré avec des techniques différentes. Il faut donc logiquement s'attendre à ce que les deux variables soient très corrélées et que les points représentant ces variables soient très proches d'une droite. Toutefois une bonne corrélation n'est pas la preuve d'une bonne concordance. Une différence systématique peut intervenir, non prise en compte dans l'évaluation de la concordance. Bae et al. avaient identifié avec leur technique originale de segmentation une sous-estimation systématique de 9% dans la mesure de leurs volumes, avec des limites d'agrément à -43.3;17.1%. Dans notre étude sur fantôme, en faisant également la transformation logarithmique, la meilleure technique amenait une sous-estimation systématique de 1% dans la mesure des volumes, avec des limites d'agrément à -3.5; 1.2 %.

Dans une étude où ils avaient testé la volumétrie à main levée sur le programme AW Server, comme dans notre étude, Kistler et al. avaient trouvé une excellente reproductibilité, avec un coefficient de corrélation de Lin de 1.000 en intra-observateur et de 0.996 en inter-observateurs (19). Dans une autre étude, Mignani et al. rapportaient une variabilité intra-observateur de 3.3% et une variabilité inter-observateurs de 4.3% (20). Bae et al. décrivaient de leur côté un coefficient de corrélation intra-observateur de 0.99 et inter-observateurs de 0.99 (12), et dans une autre étude une variabilité intra-observateur moyenne de 1.3% (16). Cohen et al. observaient une médiane d'accord intra-observateur entre 98.1% et 98.7%, et une médiane d'accord inter-observateurs entre 95.2% et 95.8% (21).

En se basant sur les corrélations de Pearson ou de Lin, les résultats des mesures cliniques de notre étude étaient donc comparables avec ceux de la littérature, avec une reproductibilité intra-observateur excellente (99.6% à 99.9% d'accord) et une concordance inter-observateurs de 99.68%. Notre étude analysait en plus les mesures de volume selon l'approche de Bland-Altman. Cette approche a montré que tant pour les mesures itératives d'un même volume que pour les mesures de volumes différents, un biais systématique de sous-évaluation s'introduisait dans la mesure. Une limite d'agrément supérieure de 90 millilitres était identifiée, ce qui signifiait qu'une augmentation apparente de moins de 90 millilitres entre deux évaluations successives de volume d'un rein ne pouvait pas être considérée comme réelle. En séparant les reins de plus ou moins 600 millilitres, ce seuil de validité baissait à 42 millilitres en dessous de 600 millilitres et montait à 120 millilitres au-dessus.

Dans l'étude de la cohorte du CRISP, Grantham et al. (8) avaient identifié qu'une croissance annuelle de 5.3% du volume rénal était associée à une baisse significative de la fonction rénale. Pour des reins de moins de 600 millilitres, ce seuil représente au plus 32 millilitres, soit bien au-dessous de la limite d'agrément de notre étude. De même, lorsque Kistler et al. (19) identifiaient une variation de  $\pm 25.3$  millilitres, cette valeur reste au-dessous de la limite d'agrément identifiée dans notre étude. L'analyse de Bland-Altman dépend des données analysées et les valeurs de biais que nous avons déterminées ainsi que les écarts de limites d'agrément ne sont pas nécessairement exploitables dans d'autres études. Sardanelli (22) a souligné l'utilité des analyses de Bland-Altman dans l'évaluation des mesures en radiologie. Cette approche permet d'identifier clairement les biais de mesure et met en lumière la plus petite différence observable. En dessous du seuil représenté par les limites d'agrément, une différence de mesure ne peut pas être considérée comme significative.

Dans l'évaluation de l'effet d'un traitement au Tolvaptan, Torres et Chapman (23) avaient identifié une différence de la croissance des reins entre les patients sous traitement et le groupe placebo. La

taille moyenne des reins évalués dans leur étude était de  $1705 \pm 921$  millilitres pour le groupe traité et  $1698 \pm 873$  millilitres pour le groupe placebo. Les modifications annuelles mesurées (respectivement  $2.8\% = 47$  millilitres et  $5.5\% = 92$  millilitres), en dessous et à la limite du seuil identifié dans notre étude, soulignent l'importance de la détermination précise de ce seuil pour obtenir une interprétation correcte des données.

Les ambiguïtés de segmentation provoquées par la présence de kystes hépatiques au contact du rein droit n'avaient pas modifié les corrélations mesurées au niveau des reins droit ou gauche. Cette homogénéité était probablement due à l'utilisation d'une technique de seuillage manuelle avec laquelle l'opérateur détermine visuellement les limites entre les kystes du rein et ceux du foie. Une technique automatique de segmentation aurait probablement confondu les deux organes.

Notre étude rétrospective a plusieurs limitations. La plus importante concerne le manque d'homogénéité du groupe des patients analysés. Lors des mesures initiales en clinique, réalisées par de multiples opérateurs peu entraînés, la technique de contour à main levée du logiciel AW avait été utilisée dans chaque cas sur des images en pondération HASTE. Certains patients avaient été examinés à 1.5 Tesla et d'autres à 3 Tesla. Pour certains patients, les coupes coronales avaient été employées pour les mesures, pour d'autres, les coupes axiales. Sur la base de l'étude du fantôme, il apparaissait qu'avec la technique de contour à main levée du logiciel AW, les volumes étaient légèrement sous-évalués en coupes axiales, tant à 1.5 Tesla qu'à 3 Tesla et qu'ils étaient surévalués en coupes coronales. Pour les coupes axiales, l'utilisation de deux champs magnétiques différents n'avait pas d'influence sur nos résultats, les mêmes images étant utilisées pour les mesures cliniques et celles de l'étude. Pour les coupes coronales, tant à 1.5 Tesla qu'à 3 Tesla, les mesures cliniques étaient surévaluées par rapport aux mesures de l'étude. La différence systématique observée dans l'étude (biais = 8.84 millilitres) et la limite supérieure d'agrément sont donc sous-estimées par rapport à la réalité. Cette limite d'agrément de 90 millilitres est donc probablement trop basse. Cette limitation souligne l'importance de disposer d'un protocole unique pour réaliser les images et les mesures itératives chez les patients. Avec les changements rapides dans le parc des machines, il faudrait même proposer de répéter les mesures précédentes lors de chaque nouvelle évaluation d'un patient souffrant d'ADPKD. Cette mesure permettrait également de compenser la rotation des internes dans les postes de formation, ce tournus ne leur permettant pas d'acquérir l'entraînement suffisant pour devenir reproductibles (21).

Nous n'avons pas évalué le temps requis pour réaliser les mesures des volumes. Plusieurs auteurs se sont intéressés à l'évaluation de techniques de segmentation automatique, destinées à réduire la durée du post-processing (17, 18, 24). Dans notre étude, certaines techniques de segmentation rendaient des résultats extrêmement reproductibles, mais entachés de biais considérables. La reproductibilité d'une mesure n'est pas garante de sa validité (25). Dans notre étude, la technique de mesure la plus fiable était une technique de segmentation manuelle, c'est-à-dire une technique demandant beaucoup de temps. Une technique de segmentation automatique fiable est donc attendue avec impatience.

En conclusion, la mesure du volume rénal par IRM montre une grande variabilité en fonction des procédés utilisés pour effectuer ces mesures. Dans notre pratique, une technique de seuillage manuel sur les coupes axiales en pondération T2 apporte les meilleurs résultats. Les mesures séquentielles des volumes rénaux au cours du temps, en employant à chaque fois la même technique de mesure, permettent d'évaluer de manière fiable les modifications de la taille des reins. Des

variations inférieures à 90 millilitres ne peuvent pas être considérées comme significatives. Au contraire, des augmentations supérieures à 90 millilitres ou des régressions de plus de 70 millilitres peuvent être interprétées comme réelles de façon fiable avec la volumétrie IRM.

## BIBLIOGRAPHIE

1. Bonny O, Chehade H, Fellmann F, Qanadli SD, Barbey F. [Recent insights into polycystic kidney disease]. *Revue medicale suisse*. 2010;6(238):454-9.
2. Bae KT, Zhu F, Chapman AB, Torres VE, Grantham JJ, Guay-Woodford LM, et al. Magnetic resonance imaging evaluation of hepatic cysts in early autosomal-dominant polycystic kidney disease: the Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease cohort. *Clinical journal of the American Society of Nephrology : CJASN*. 2006;1(1):64-9.
3. Torra R, Sarquella J, Calabia J, Marti J, Ars E, Fernandez-Llama P, et al. Prevalence of cysts in seminal tract and abnormal semen parameters in patients with autosomal dominant polycystic kidney disease. *Clinical journal of the American Society of Nephrology : CJASN*. 2008;3(3):790-3.
4. Choukroun G, Itakura Y, Albouze G, Christophe JL, Man NK, Grunfeld JP, et al. Factors influencing progression of renal failure in autosomal dominant polycystic kidney disease. *Journal of the American Society of Nephrology : JASN*. 1995;6(6):1634-42.
5. Gabow PA. Autosomal dominant polycystic kidney disease. *The New England journal of medicine*. 1993;329(5):332-42.
6. Martin PY. L'insuffisance rénale chronique. *Le Fait Médical*. 2009;Déc.
7. Grantham JJ, Chapman AB, Torres VE. Volume progression in autosomal dominant polycystic kidney disease: the major factor determining clinical outcomes. *Clinical journal of the American Society of Nephrology : CJASN*. 2006;1(1):148-57.
8. Grantham JJ, Torres VE, Chapman AB, Guay-Woodford LM, Bae KT, King BF, Jr., et al. Volume progression in polycystic kidney disease. *The New England journal of medicine*. 2006;354(20):2122-30.
9. Pei Y, Hwang YH, Conklin J, Sundsbak JL, Heyer CM, Chan W, et al. Imaging-Based Diagnosis of Autosomal Dominant Polycystic Kidney Disease. *Journal of the American Society of Nephrology : JASN*. 2014.
10. Ravine D, Gibson RN, Walker RG, Sheffield LJ, Kincaid-Smith P, Danks DM. Evaluation of ultrasonographic diagnostic criteria for autosomal dominant polycystic kidney disease 1. *Lancet*. 1994;343(8901):824-7.
11. Chapman AB, Wei W. Imaging approaches to patients with polycystic kidney disease. *Seminars in nephrology*. 2011;31(3):237-44.
12. Bae KT, Tao C, Zhu F, Bost JE, Chapman AB, Grantham JJ, et al. MRI-based kidney volume measurements in ADPKD: reliability and effect of gadolinium enhancement. *Clinical journal of the American Society of Nephrology : CJASN*. 2009;4(4):719-25.
13. Thomsen HS, European Society of Urogenital R. ESUR guideline: gadolinium-based contrast media and nephrogenic systemic fibrosis. *European radiology*. 2007;17(10):2692-6.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-10.
15. Serra AL, Wüthrich RP. Maladie polykystique rénale autosomale dominante (ADPKD): pathogénèse, clinique et nouvelles options thérapeutiques. *Forum Med Suisse*. 2009;9(47):854-8.
16. Bae KT, Commean PK, Lee J. Volumetric measurement of renal cysts and parenchyma using MRI: phantoms and patients with polycystic kidney disease. *Journal of computer assisted tomography*. 2000;24(4):614-9.

17. Bae KT, Sun H, Lee JG, Bae K, Wang J, Tao C, et al. Novel methodology to evaluate renal cysts in polycystic kidney disease. *American journal of nephrology*. 2014;39(3):210-7.
18. Bae KT, Tao C, Wang J, Kaya D, Wu Z, Bae JT, et al. Novel approach to estimate kidney and cyst volumes using mid-slice magnetic resonance images in polycystic kidney disease. *American journal of nephrology*. 2013;38(4):333-41.
19. Kistler AD, Poster D, Krauer F, Weishaupt D, Raina S, Senn O, et al. Increases in kidney volume in autosomal dominant polycystic kidney disease can be detected within 6 months. *Kidney international*. 2009;75(2):235-41.
20. Mignani R, Corsi C, De Marco M, Caiani EG, Santucci G, Cavagna E, et al. Assessment of kidney volume in polycystic kidney disease using magnetic resonance imaging without contrast medium. *American journal of nephrology*. 2011;33(2):176-84.
21. Cohen BA, Barash I, Kim DC, Sanger MD, Babb JS, Chandarana H. Intraobserver and interobserver variability of renal volume measurements in polycystic kidney disease using a semiautomated MR segmentation algorithm. *AJR American journal of roentgenology*. 2012;199(2):387-93.
22. Sardanelli F, Di Leo G. Biostatistics for Radiologists. In: Springer, editor. Milan: Springer-Verlag; 2009. p. 125-40.
23. Torres VE, Chapman AB, Devuyst O, Gansevoort RT, Grantham JJ, Higashihara E, et al. Tolvaptan in patients with autosomal dominant polycystic kidney disease. *The New England journal of medicine*. 2012;367(25):2407-18.
24. Bae K, Park B, Sun H, Wang J, Tao C, Chapman AB, et al. Segmentation of individual renal cysts from MR images in patients with autosomal dominant polycystic kidney disease. *Clinical journal of the American Society of Nephrology : CJASN*. 2013;8(7):1089-97.
25. Kopans DB. Basic physics and doubts about relationship between mammographically determined tissue density and breast cancer risk. *Radiology*. 2008;246(2):348-53.

## ANNEXES

**Tableau 1.** Pesée des 8 éléments avant et après assemblage

	Mesure 1	Mesure 2	Mesure 3	Mesure 4	Mesure 5	Moyenne
Masse des kystes (g)	715.15	715.10	715.08	715.09	715.09	<b>715.1</b>
Masse du fantôme (g)	717.40	717.39	717.39	717.41	717.40	<b>717.4</b>

**Tableau 2.** Volume du fantôme obtenu par déplacement des fluides.\*effectuée sur les mesures 2 à 4. La mesure 1, effectuée avant que le système n'atteigne son état d'équilibre, n'avait pas été prise en compte dans les calculs de volume.

	Mesure 1	Mesure 2	Mesure 3	Mesure 4	Moyenne*
Volume fantôme	675 ml	725 ml	710 ml	710 ml	<b>715 ml</b>

**Tableau 3.** Analyse de Bland-Altman des mesures de volume réalisées avec les différents programmes sur les deux machines (Aera 1.5T et Verio 3T) et les différents plans. Bias correspond à la différence moyenne des mesures, SD à la déviation standard des différences, limit of agreement (LoA) correspond au biais  $\pm 2SD$ . La technique Aera T2 trans AW main avait le meilleur rapport entre le biais et les LoA. La technique Aera T2 cor CS auto avait le plus petit biais, au dépend d'une très grande dispersion.

	Bias (ml)	SD (ml)	Lower limit of agreement (ml)	Upper limit of agreement (ml)
Aera T2 trans CS main	19.49	7.71	4.07	34.90
Aera T2 trans CS auto	-56.52	18.69	-93.90	-19.14
<b>Aera T2 trans AW main</b>	<b>-8.15</b>	<b>8.52</b>	<b>-25.19</b>	<b>8.88</b>
Aera T2 trans AW auto	-60.95	2.80	-66.55	-55.35
Aera T2 trans AW seuil	-40.70	5.13	-50.97	-30.44
Aera T2 cor CS main	62.77	19.41	23.96	101.59
<b>Aera T2 cor CS auto</b>	<b>-3.35</b>	<b>52.13</b>	<b>-107.61</b>	<b>100.91</b>
Aera T2 cor AW main	55.27	14.27	26.74	83.81
Aera T2 cor AW auto	-46.05	5.48	-57.01	-35.09
Aera T2 cor AW seuil	-25.21	19.05	-63.30	12.89
Aera T1 cor CS main	61.52	13.86	33.79	89.25
Aera T1 cor CS auto	7.06	22.44	-37.82	51.94
Aera T1 cor AW main	46.01	9.68	26.66	65.37
Verio T2 trans CS main	18.84	9.05	0.73	36.95
Verio T2 trans CS auto	-95.70	19.39	-134.47	-56.92
Verio T2 trans AW main	-25.21	10.14	-45.49	-4.92
Verio T2 trans AW auto	-82.25	4.25	-90.76	-73.74
Verio T2 trans AW seuil	-68.65	4.63	-77.91	-59.40
Verio T2 cor CS main	82.89	13.00	56.89	108.88
Verio T2 cor CS auto	-38.69	70.99	-180.68	103.29
Verio T2 cor AW main	29.36	7.79	13.79	44.94
Verio T2 cor AW auto	-42.65	2.78	-48.20	-37.10
Verio T2 cor AW seuil	-28.71	4.00	-36.71	-20.71
Verio T1 cor CS main	78.03	13.87	50.28	105.77
Verio T1 cor CS auto	-29.68	18.36	-66.40	7.05
Verio T1 cor AW main	13.82	11.22	-8.62	36.26

**Tableau 4.** Comparaison des résultats suivant la machine IRM utilisée. (1) calculé comme étant  $100 * ((\text{Aera} - \text{Verio}) / \text{Aera})$ .

	Aera	Verio	% d'écart (1)
CS main T2 transverse	734.487	733.842	0.1
CS auto T2 transverse	658.483	619.304	5.9
AW main T2 transverse	706.848	689.793	2.4
AW contour T2 transverse	654.05	632.75	3.3
AW seuil T2 transverse	674.297	646.346	4.1
CS main T2 coronal	777.775	797.887	-2.6
CS auto T2 coronal	711.652	676.307	5
AW main T2 coronal	770.273	744.364	3.4
AW contour T2 coronal	668.95	672.35	-0.5
AW seuil T2 coronal	689.794	686.287	0.5
CS main T1 coronal	776.522	793.027	-2.1
CS auto T1 coronal	722.064	685.324	5.1
AW main T1 coronal	761.014	728.823	4.2

**Tableau 5.** Reproductibilité intra-observateur. La première colonne de chaque patient contient les volumes du rein gauche, la seconde ceux du rein droit.

Patient	A	A	B	B	C	C	D	D	E	E
Mesure 1 (ml)	418	587	1458	1574	940	1083	326	226	876	659
Mesure 2 (ml)	404	585	1484	1571	935	1076	323	224	872	650
Mesure 3 (ml)	401	583	1485	1579	936	1072	317	222	874	659
Moyenne (ml)	407	585	1476	1575	937	1077	322	224	874	656
<b>Concordance (%)</b>	<b>98.3</b>	<b>99.8</b>	<b>99.2</b>	<b>99.8</b>	<b>99.8</b>	<b>99.6</b>	<b>99.0</b>	<b>99.4</b>	<b>99.9</b>	<b>99.4</b>

**Tableau 6.** Indices de corrélation inter-observateurs. Le coefficient de Pearson indique la précision, le coefficient de Lin l'exactitude (accuracy).

<b>Indices de corrélation inter-observateurs</b>	
Moyenne (%)	95.2
Médiane (%)	96.1
CV	0.04
Pearson	0.997
Lin	0.997

**Tableau 7.** Indices de concordance inter-observateurs rein gauche-rein droit, et petit volume-grand volume. Le coefficient de Pearson indique la précision, le coefficient de Lin l'exactitude (accuracy).

	<b>Rein gauche</b>	<b>Rein droit</b>
<b>Moyenne (%)</b>	94.8	95.6
<b>Médiane (%)</b>	95.8	96.1
<b>CV</b>	0.04	0.04
<b>Pearson</b>	0.997	0.997
<b>Lin</b>	0.997	0.997

	<b>&lt;600 ml</b>	<b>&gt;600 ml</b>
<b>Moyenne (%)</b>	95.0	95.4
<b>Médiane (%)</b>	96.2	95.6
<b>CV</b>	0.05	0.03
<b>Pearson</b>	0.987	0.995
<b>Lin</b>	0.987	0.994

## Légendes des figures

**Figure 1.** A. Les éléments de polystyrène avant le montage. B. Les éléments montés sous forme de rein polykystique.

**Figure 2.** A. Le dispositif destiné à la mesure du volume par la méthode de déplacement des fluides. B. Le fantôme de rein polykystique juste avant d'être immergé dans le bac de mesure.

**Figure 3.** Le modèle de rein polykystique immergé dans le bac de saindoux solidifié

**Figure 4.** A. Une des machines IRM utilisée (Verio 3T). B. Le fantôme devant le tunnel de l'IRM (Aera 1.5T)

**Figure 5.** A. Image transverse HASTE T2 obtenue avec la machine Aera 1.5T. B. Image coronale HASTE T2 obtenue avec la machine Verio 3T.

**Figure 6.** Analyse de volume sur le logiciel AW Server avec des coupes transverses HASTE T2.

**Figure 7.** A. Rendu de volume d'un rein droit d'un volume de 1106.96 millilitres. B. Rendu de volume d'un rein gauche d'un volume de 1262.15 millilitres.

**Figure 8.** Box-plot des mesures du fantôme sur les séries acquises à 1.5T. La ligne au milieu de la boîte représente la médiane (50<sup>ème</sup> percentile). La boîte s'étend du 25<sup>ème</sup> ( $x_{(25)}$ ) au 75<sup>ème</sup> ( $x_{(75)}$ ) percentile et représente la fourchette interpercentile (IQR). Les lignes prolongeant les boîtes vers le bas ou vers le haut représentent les valeurs inférieures et supérieures encore dans les limites définies par  $(x_{(25)}) - 1.5 \times \text{IQR}$  et  $(x_{(75)}) + 1.5 \times \text{IQR}$ . Les valeurs plus extrêmes observées sont représentées individuellement. La technique la plus reproductible (T2 trans AW auto) sous-estimait considérablement le volume. La technique qui donnait le meilleur résultat moyen (T2 cor CS auto) avait par contre une grande dispersion des résultats.

**Figure 9.** Box-plot des mesures du fantôme sur les séries acquises à 3T. La technique la plus reproductible (T2 cor AW auto) sous-estimait le volume. La technique qui donnait le meilleur résultat moyen (T2 cor AW main) avait par contre une importante dispersion des résultats.

**Figure 10.** Fiabilité des différentes méthodes de volumétrie.

**Figure 11.** Graphique de Bland-Altman de la variabilité intra-observateur. Les mesures sont systématiquement surestimées avec un biais de 7.79. Les limits of agreement sont à -71.20;86.78.

**Figure 12.** Graphique de Bland-Altman de la variabilité intra-observateur, en tenant compte d'un seuil de volume de 600 millilitres. Les mesures sont systématiquement surestimées avec un biais de 11.85 millilitres pour les reins <600ml et 5.09 millilitres pour les reins >600 millilitres. La LoA est de -18.25 ;42.05 pour les reins <600ml et -94.79 ;104.97 pour les reins >600ml.

**Figure 13.** Diagramme de dispersion avec droite de régression, comparant les mesures d'origine avec les mesures effectuées dans le cadre de cette étude

**Figure 14.** Graphique de Bland-Altman de la variabilité inter-observateur. Les mesures sont systématiquement surestimées avec un biais de 8.84 ( $\sigma=4.71$ , -0.56 ;18.24). Les LoA sont à -72.32 ;90.00.

**Figure 15.** Graphique de Bland-Altman de la variabilité inter-observateur en séparant les reins <600ml et >600ml. Pour les reins <600ml, les mesures sont justes avec un biais de 0.7. Les LoA sont à -41.44 ;42.86. Pour les reins >600ml, les mesures sont surévaluées avec un biais de 16.55. Les LoA sont à -85.04 ;118.13.

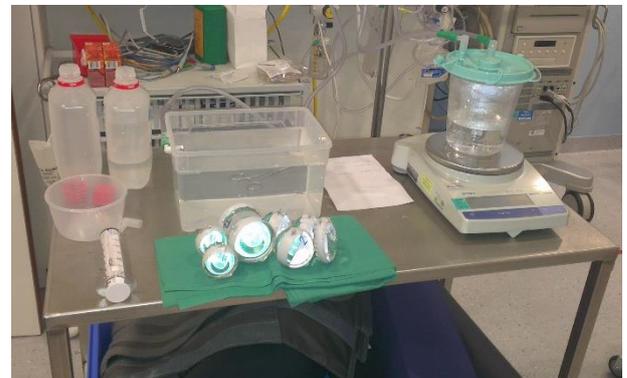
**Figure 1**



A

B

**Figure 2**



A

B

**Figure 3**



**Figure 4**

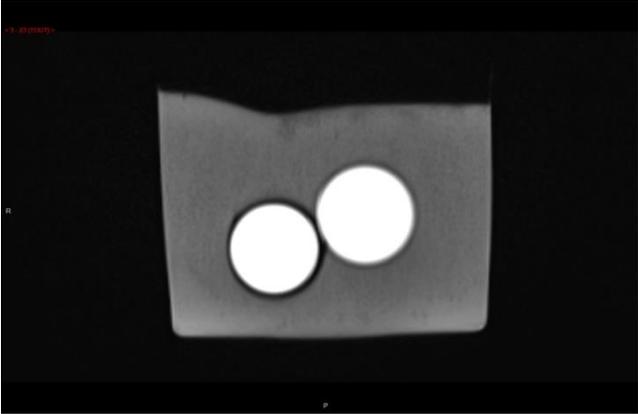


A

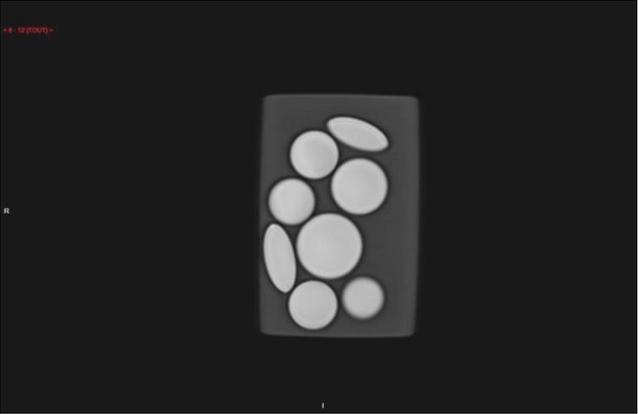


B

**Figure 5**



A

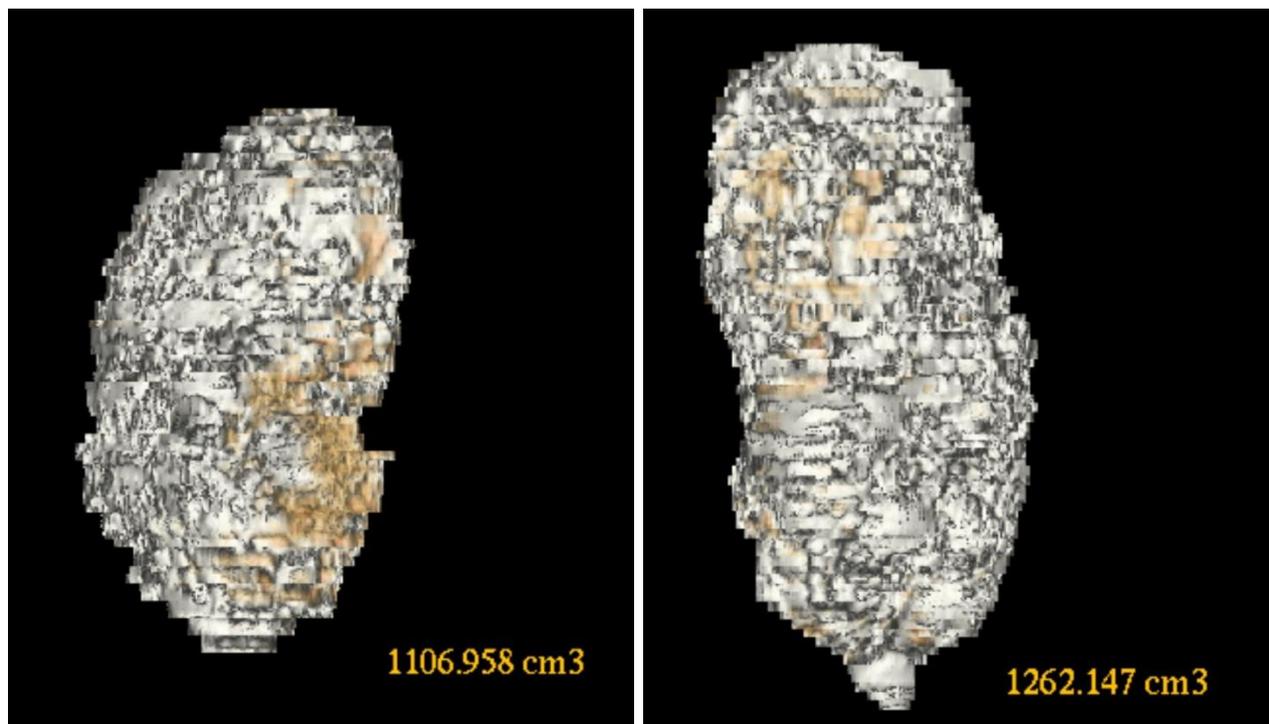


B

**Figure 6**



**Figure 7**



A

B

Figure 8

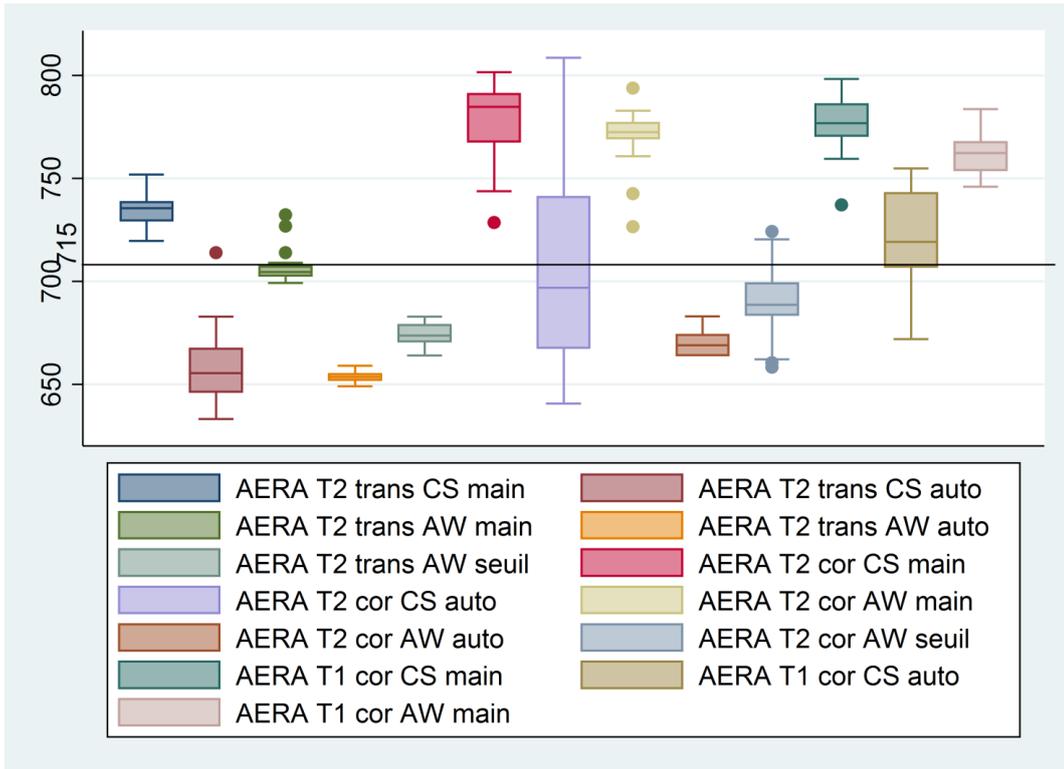
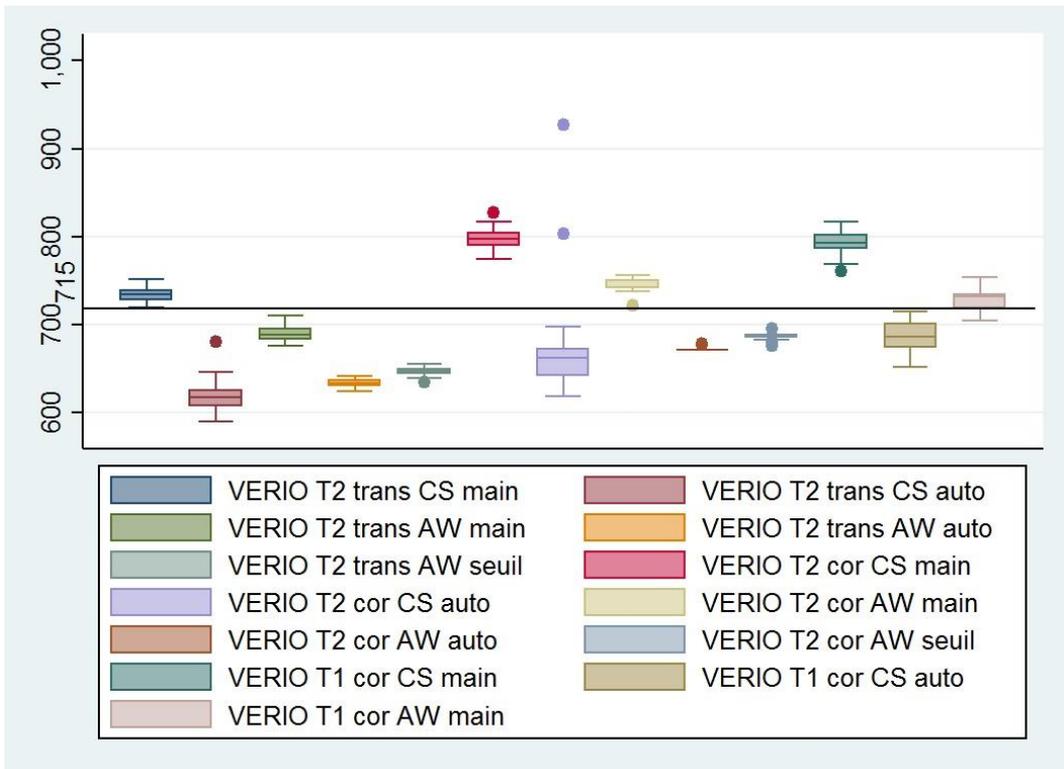
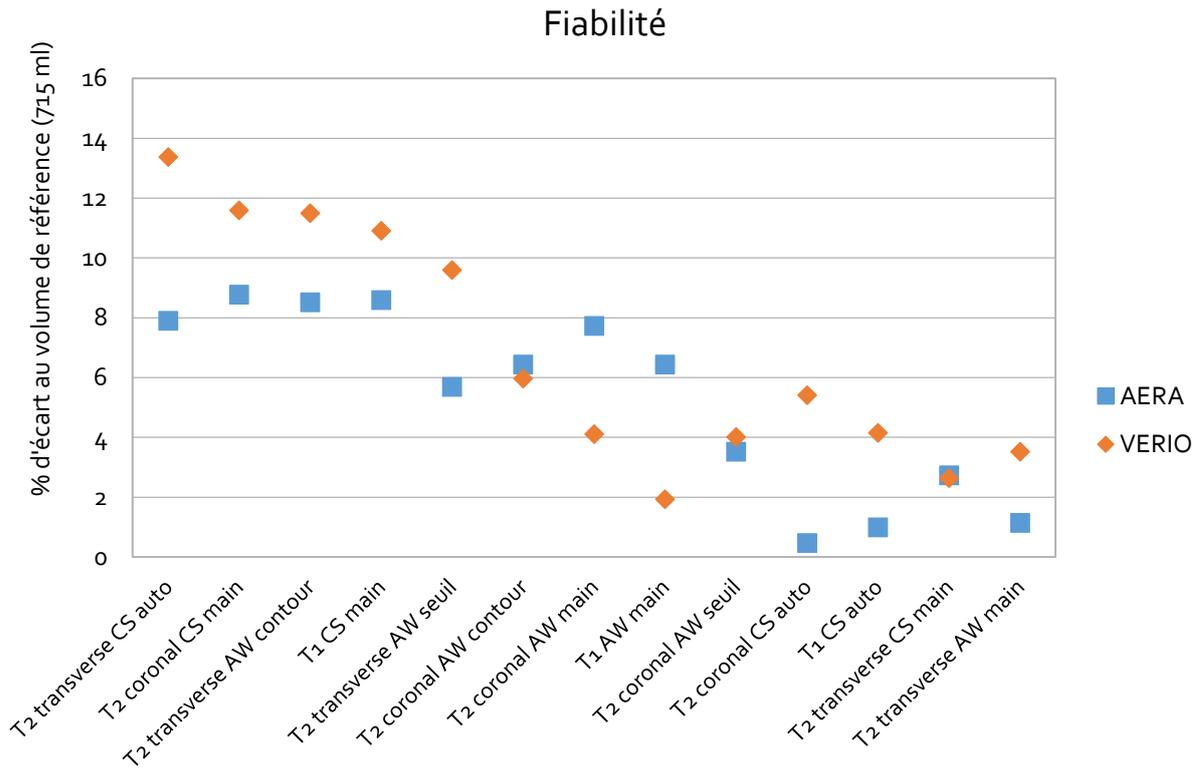


Figure 9



**Figure 10**



**Figure 11**

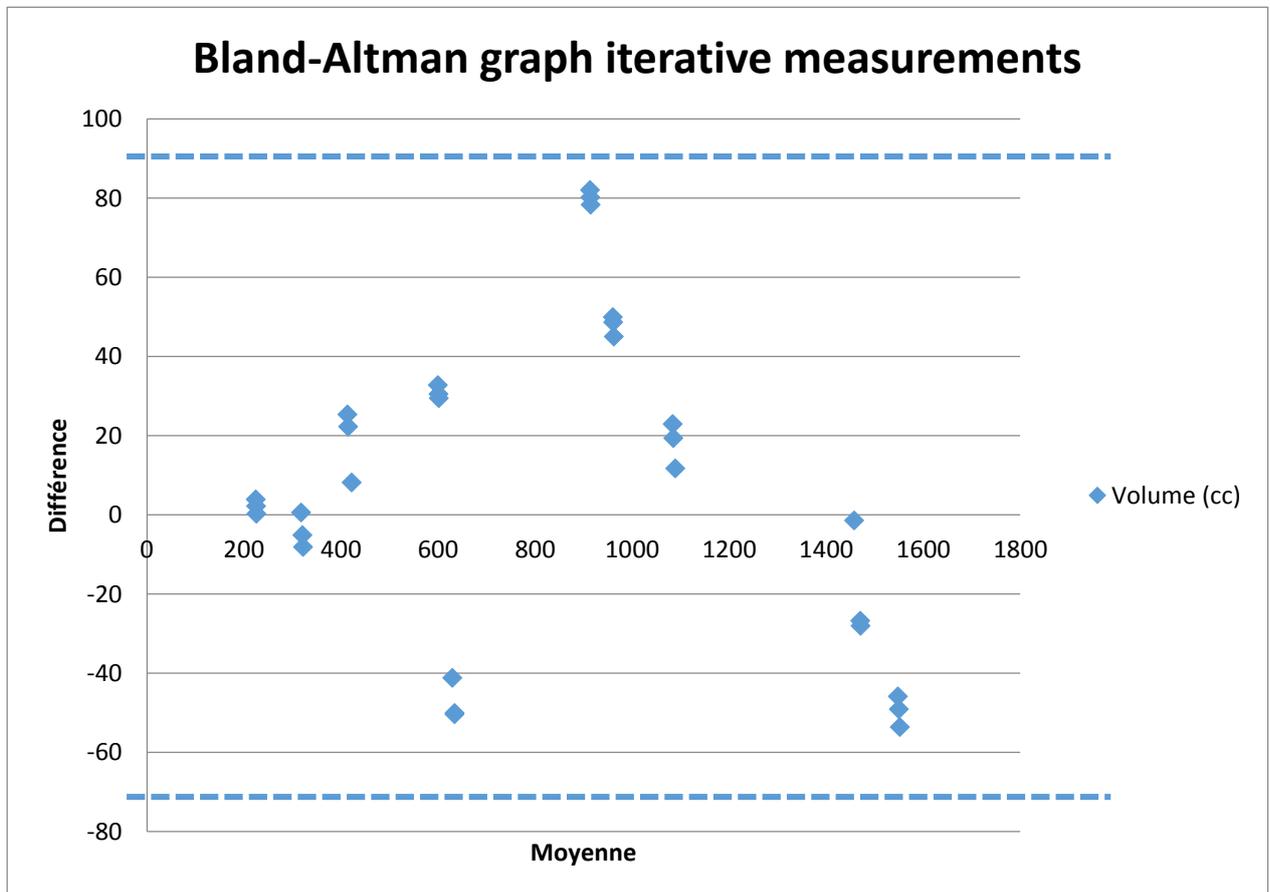


Figure 12

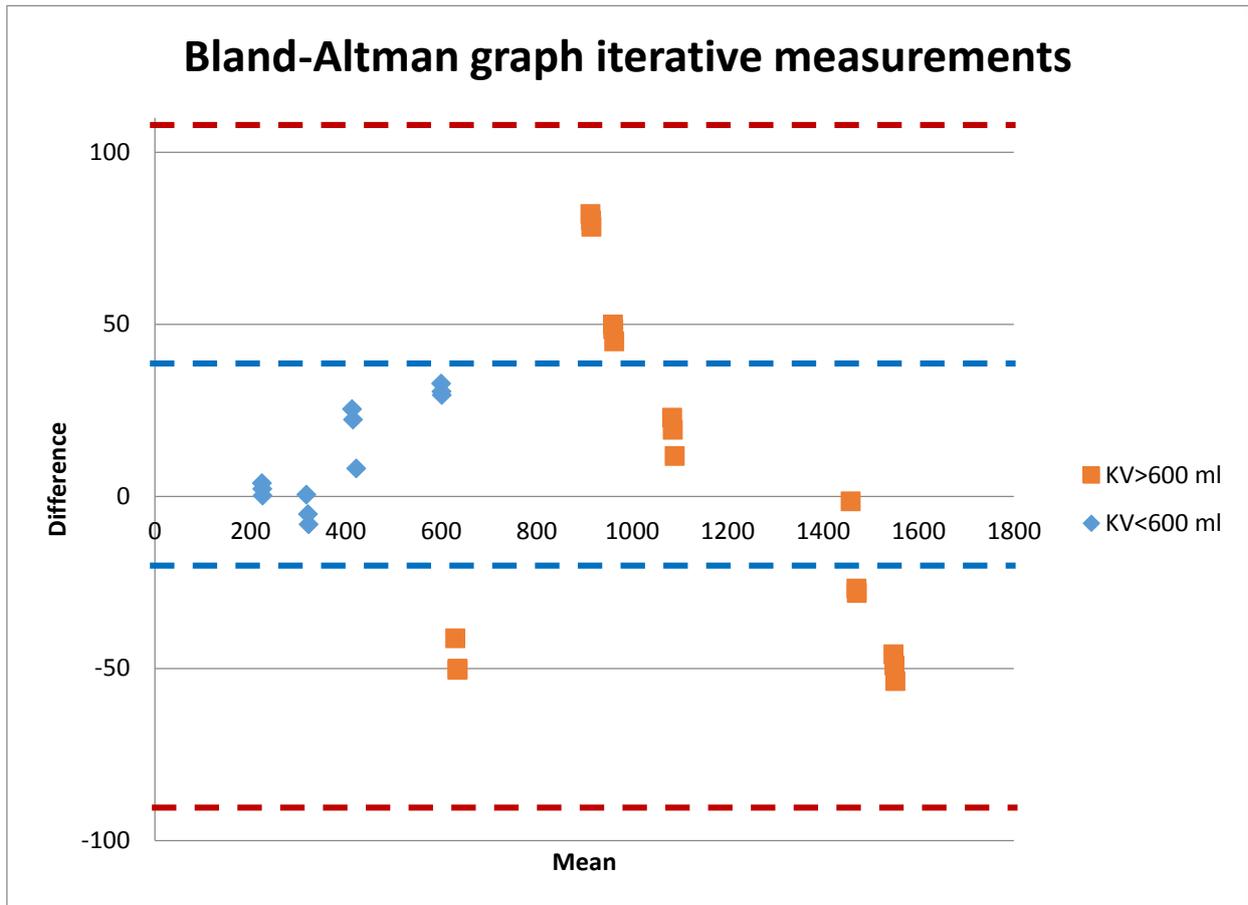


Figure 13

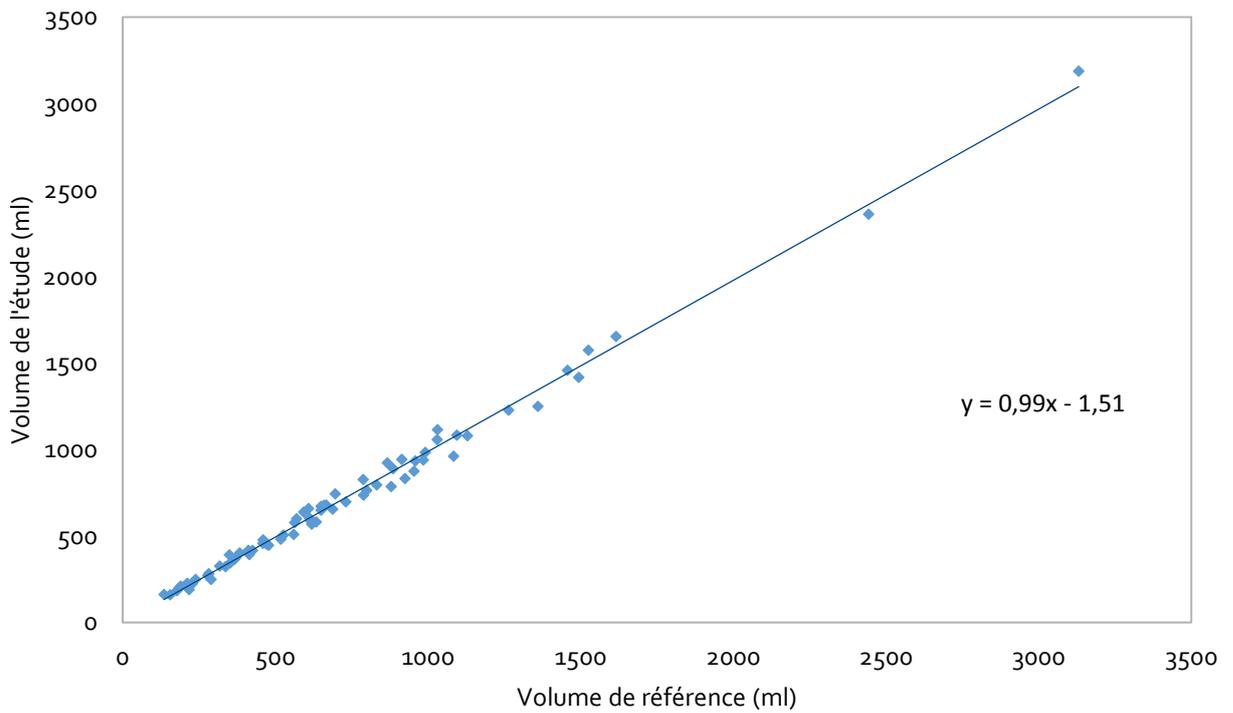


Figure 14

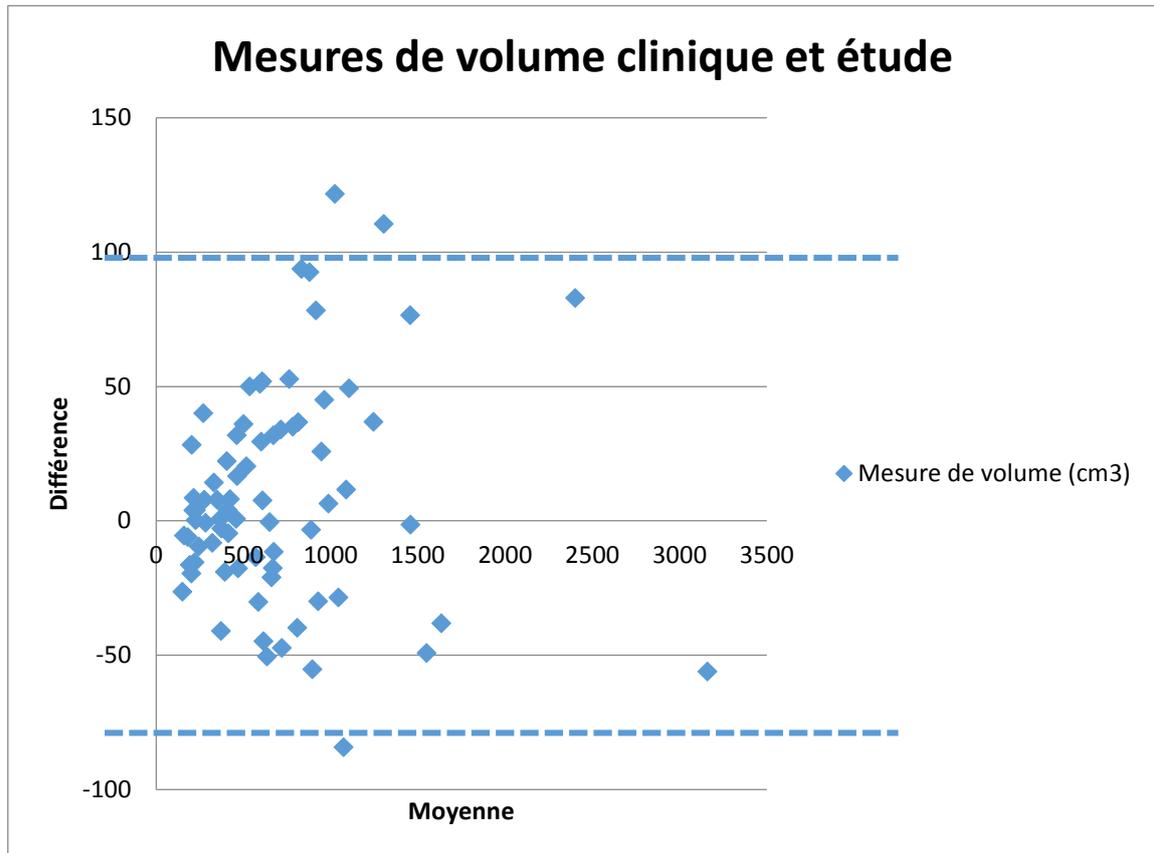


Figure 15

