Measurements in Idiopathic Normal Pressure Hydrocephalus
Computerized neuropsychological test battery and intracranial pulse waves

Anders Behrens

Department of Pharmacology and Clinical Neuroscience
Department of Radiation Sciences, Umeå University
Umeå 2014
"Extraordinary claims require extraordinary evidence."

Carl Sagan

"Disorders of intellect [...] happen much more often than superficial observers will easily believe."

Samuel Johnson: The History of Rasselas, Prince of Abyssinia (1759)

"Nog finns det mål och mening i vår färd. Men det är vägen som är mödan värd."

Karin Boye

To my family Fanny, Astrid and Lovisa
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE OF CONTENTS</td>
<td>ii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>ORIGINAL PAPERS</td>
<td>vi</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td>Non-invasive intracranial pressure</td>
<td>1</td>
</tr>
<tr>
<td>Transcranial Doppler ultrasonography</td>
<td>2</td>
</tr>
<tr>
<td>ICP pulsations</td>
<td>4</td>
</tr>
<tr>
<td>Idiopathic Normal Pressure Hydrocephalus</td>
<td>6</td>
</tr>
<tr>
<td>Hydrocephalus overview</td>
<td>6</td>
</tr>
<tr>
<td>Idiopathic Normal Pressure Hydrocephalus</td>
<td>7</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>7</td>
</tr>
<tr>
<td>Pathophysiology of INPH</td>
<td>8</td>
</tr>
<tr>
<td>Diagnosing INPH</td>
<td>9</td>
</tr>
<tr>
<td>Treatment of INPH</td>
<td>9</td>
</tr>
<tr>
<td>Supplemental tests</td>
<td>10</td>
</tr>
<tr>
<td>Cognitive testing in INPH</td>
<td>11</td>
</tr>
<tr>
<td>Neuropsychological testing in INPH</td>
<td>12</td>
</tr>
<tr>
<td>Computerized neuropsychological testing</td>
<td>15</td>
</tr>
<tr>
<td>Test development</td>
<td>16</td>
</tr>
<tr>
<td>Research rationale</td>
<td>16</td>
</tr>
<tr>
<td>AIMS</td>
<td>18</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td></td>
</tr>
<tr>
<td>Ethical approval</td>
<td>19</td>
</tr>
<tr>
<td>Patients and controls</td>
<td></td>
</tr>
<tr>
<td>INPH patients of paper I and II</td>
<td>19</td>
</tr>
<tr>
<td>Cognitively impaired patients of paper III</td>
<td>20</td>
</tr>
<tr>
<td>Hydrocephalus patients of paper III</td>
<td>20</td>
</tr>
<tr>
<td>Healthy elderly of papers III and IV</td>
<td>20</td>
</tr>
<tr>
<td>INPH patients of paper IV</td>
<td>21</td>
</tr>
<tr>
<td>Simultaneous infusion test, transcranial Doppler and pressure measurements</td>
<td>22</td>
</tr>
<tr>
<td>Mathematical model</td>
<td>24</td>
</tr>
<tr>
<td>Bench test</td>
<td>25</td>
</tr>
<tr>
<td>Neuropsychological test battery</td>
<td>26</td>
</tr>
<tr>
<td>Software design</td>
<td>26</td>
</tr>
<tr>
<td>Computerized General Neuropsychological INPH Test (CoGNIT)</td>
<td>26</td>
</tr>
<tr>
<td>Statistics</td>
<td>29</td>
</tr>
<tr>
<td>RESULTS</td>
<td>31</td>
</tr>
<tr>
<td>Transcranial Doppler pulsatility index and intracranial pressure</td>
<td>31</td>
</tr>
<tr>
<td>Intracranial and lumbar pressure waves</td>
<td>32</td>
</tr>
<tr>
<td>Baseline</td>
<td>33</td>
</tr>
</tbody>
</table>
ABSTRACT

Idiopathic Normal Pressure Hydrocephalus (INPH) is a condition affecting gait, cognition and continence. Radiological examination reveals enlarged ventricles of the brain. The pathophysiology is debated, but the disease is probably related to abnormal pressures and flows in the Cerebrospinal fluid (CSF) dynamic system. A shunt that drains CSF from the ventricles to the abdomen often improves the symptoms. Much research on INPH has been focused on identifying tests that distinguish INPH from other disease groups, and predict the outcome after shunt surgery. As part of these quests, there are attempts to find measurement methods of intracranial parameters that are valid, reliable, tolerable and safe for patients.

Today's technologies for intracranial pressure (ICP) measurement are invasive, often requiring a burr-hole in the skull. Recently, a method for non-invasive ICP measurements was suggested: the Pulsatile Index (PI) calculated from transcranial Doppler data assessed from the middle cerebral artery. In previous studies, the PI is strongly correlated with ICP. The PI method has been proposed as an initial screening tool for neurointensive care patients, aiding decisions of who would need additional invasive pressure monitoring. However, there is great discrepancy between studies in the suggested calibration of PI and ICP data. In paper I, the relation between PI and ICP was explored in INPH patients during controlled ICP regulation by lumbar infusion. The confidence interval for predicted ICP, based on measured PI was too large for the method to be of clinical utility.

There is a quest for better predictive tests for shunt success in INPH. Recent studies show promising results with criteria based on ICP wave amplitudes. The brain ventricular system, and the fluid surrounding the spinal cord are in contact. In paper II, it was shown that ICP waves could be measured via lumbar subarachnoid space, with a slight underestimation. Less invasive than traditional methods of measuring intracranial pressure, this technique would be associated with fewer complications.

One of the cardinal symptoms of hydrocephalus is cognitive impairment. Neuropsychological studies have demonstrated specific cognitive domains that are impaired and improve after shunt surgery in INPH patients. These include memory, psychomotor speed, attention, manual dexterity and executive functions. However, there is currently no standardized test battery, and different studies use different tests and outcome measures. Therefore it is often difficult to compare results from different research groups. In response, in this thesis a fully automated computerized neuropsychological
test battery was developed. Included tests were selected based on a literature search about evidence for neuropsychological testing in INPH. In papers III and IV the validity, test-retest reliability, responsiveness to improvement after shunt surgery and feasibility for testing INPH patients was demonstrated. It was also demonstrated that a group of patients with INPH were impaired in all subtests, compared to healthy elderly.
This thesis is based on the following papers, which are referred to by their Roman numerals in the text. Also, correspondence regarding paper I is included.


*Reprinted with permission*
INTRODUCTION

This thesis involves measurements of intracranial pressure (ICP), pressure pulsatility and cognitive functions in patients with idiopathic normal pressure hydrocephalus (INPH). The introduction will give the relevant physiological and technical background for understanding the disease and the methods used in the studies.

Non-invasive intracranial pressure

Measurement of ICP is of vital importance in neurointensive care (1). Secondary brain injury due to elevated ICP is an important cause of morbidity in patients with acute neurological and neurosurgical conditions such as: Traumatic Brain Injury (TBI), intracerebral hemorrhage, subarachnoid hemorrhage, hydrocephalus, malignant infarction, cerebral oedema and infections of the central nervous system (2,3). Direct measurement of ICP requires neurosurgical procedures, implanting a probe into the skull. External ventricular drainage, where a catheter is placed into one of the ventricles via a burr hole is regarded as the gold standard (2). However, the technique is associated with complications such as hemorrhages and infections (4,5).

Non-invasive measurement of ICP would have several advantages. Obviously the risk of complications caused by the neurosurgical procedure would disappear. Also, measurements could be performed in locations without access to neurosurgeons, such as smaller hospitals, ambulances and even in space (6). Several methods for non-invasive ICP assessment have been proposed. Here follows a brief and non-exhausting discussion of methods representing the principal ways of how non-invasive ICP can be assessed. The transcranial Doppler sonography method is covered in more depth.

Tympanic Membrane Displacement relies on the communication between the CSF and the perilymph via the cochlear aqueduct. The method utilizes tympanometry to measure the effect of the stapedial reflex. As stapes rests on the oval window, its resulting movement is dependent on the cochlear fluid pressure, and therefore also the CSF pressure. The resulting tympanic movement correlates with ICP. However, the predictive limits are approximately ± 25 mm Hg making this method to imprecise for clinical use (7).
Optic nerve sheath diameter measurements rely on the fact that the subarachnoid space extends around the optic nerve, meaning that changes in ICP are reflected in the diameter of the nerve sheath. Thus, ICP can be assessed via transocular ultrasound. The sensitivity and specificity to detect ICP > 20 mm Hg is within the range of 87–95% and 79–100% respectively (8). However, the nerve sheath diameter may also be affected by a number of medical conditions potentially making the method less useful (2).

Magnetic Resonance Imaging (MRI) utilizes motion-sensitive MRI to measure the blood volumes entering and leaving the cranium during one cardiac cycle and the resulting CSF-flow. From these values an elastance index can be derived, that have shown high correlation with ICP in one study (9).

Transcranial Doppler ultrasonography

Transcranial Doppler ultrasonography (TCD) was first introduced in 1982, as a way to assess blood flow velocity in the basal cerebral arteries stemming from the circle of Willis (10). An ultrasonic wave of known frequency is directed at the blood flow and the reflected signal is measured. The difference in frequency, the so-called Doppler frequency shift, is assessed and used to calculate the velocity. The most commonly assessed vessel is the middle cerebral artery (MCA), because of ease of access and good quality of the signal. The MCA carries 50-60% of ipsilateral internal carotid blood flow and is thus the largest artery emanating from the circle of Willis (11). It is insonated through the temporal bone window, found in an area spanned by a thought line from tragus to the lateral canthus of the eye, and two cm above this line.

The flow velocity is pulsatile in synchrony with the heartbeat, and from the TCD signal, systolic, diastolic and time-averaged velocities can be calculated. Normal values of mean velocity in healthy adults range between 46-86 cm/s (10). An increase in flow velocity would indicate decreased cross sectional area of the MCA, as in vasospasm after subarachnoid hemorrhage (12).

Pulsatility index (PI) is an index calculated from ultrasonic velocity waveforms. Gosling first introduced the index in 1974, as an indicator of vascular resistance. He applied PI differences to diagnose peripheral vascular disease in the popliteal arteries (13). The index is calculated as:

\[
\frac{(V_{sys} - V_{dia})}{V_{mean}}
\]
The terms $V_{\text{sys}}$, $V_{\text{dia}}$ and $V_{\text{mean}}$ are the maximum, minimum and mean flow velocities during the cardiac cycle. The main advantage of the PI is that the nominator and denominator of the ratio are equally affected by the angle of insonation, and the index is independent of this parameter. This makes for a stable index, unaffected by small inclination deviances of the ultrasonic probe.

There have been a number of studies relating PI to ICP, giving hopes for a good and accessible non-invasive ICP monitoring technique. Results are reviewed in table 1.

**Table 1.** Studies on Pulsatility Index (PI) and ICP published before paper I.

<table>
<thead>
<tr>
<th>Study</th>
<th>Clinical material</th>
<th>Number of patients</th>
<th>Calibration</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homburg et al. (14)</td>
<td>TBI</td>
<td>10</td>
<td>$\text{ICP} = 42 \times \ln \text{PI} + 14$</td>
<td>0.67</td>
</tr>
<tr>
<td>Voulgaris et al. (15)</td>
<td>TBI</td>
<td>37</td>
<td>$\text{ICP} = 16 \times \text{PI} + 5$</td>
<td>0.41</td>
</tr>
<tr>
<td>Moreno et al. (16)</td>
<td>TBI</td>
<td>125</td>
<td>$\text{ICP} = 33 \times \text{PI} - 16$</td>
<td>0.69</td>
</tr>
<tr>
<td>Bellner et al. (17)</td>
<td>TBI and SH</td>
<td>81</td>
<td>$\text{ICP} = 11 \times \text{PI} - 1$</td>
<td>0.88</td>
</tr>
</tbody>
</table>

TBI = Traumatic Brain Injury, SH = Subarachnoid Hemorrhage, ICP = Intracranial Pressure, PI = Pulsatility Index.

The most positive results are from the study by Bellner et al (17). They reported a sensitivity and specificity of 0.89 and 0.92 to detect ICP higher than 20 mm Hg. The conclusion was that "PI may be of guiding value in the invasive ICP placement decision in the neurointensive care patient".

There certainly seems to be a strong correlation between PI and ICP. However, there are large differences between regression lines of the PI and ICP relationship between studies, giving largely different estimations of ICP. Before clinical acceptance of the method, the reliability of ICP estimations needs to be examined. Acceptance of an unreliable method, especially when used in critically ill patients, may lead to harmful clinical decisions.

It is therefore important to thoroughly evaluate this method. This was the rationale for **paper I**. Using lumbar infusion of CSF, ICP could be controlled, which made it possible to obtain measurement points over a wide
range of ICPs in each patient. Thus, making it possible to study the PI/ICP relation on an individual level.

ICP pulsations

The intracranial cavity mainly consists of three largely incompressible substances: brain parenchyma, blood and CSF. As the volume of the cavity is close to constant, a volume change of any of these three substances must be compensated by a similar negative volume change of the others. This relation has been termed the Monro-Kellie doctrine (18). Marmarou et al. have modelled the intracranial response to a volume change, with later modification by Avezaat and Eijndhoven (19,20):

\[ ICP = P_1 e^{kV} + P_0 \]

The volume parameter (V) is the deviation from the equilibrium volume. The term k describes the elastic properties of the system and is called elastance and the \( P_0 \) and \( P_1 \) are pressure constants. The equation describes the so-called pressure volume curve, displayed in Figure 1. It is evident that there is an exponential increase in ICP at higher added volumes.

![Figure 1](image)

**Figure 1.** The intracranial pressure-volume curve. The same amount of volume entering the skull will produce different pressure amplitudes, depending on the position along the pressure-volume curve.

In the systole of each heartbeat a bolus of blood is leaving the heart, and transmitted through the arterial vascular tree. As described by the
Marmarou equation, the systolic arterial expansion will, when entering the scull lead to an increased ICP. The amplitude of the resulting ICP pulse is dependent on the volume of blood entering the cranium (arterial expansion), and the subject’s position along the pressure-volume curve. In Figure 1 it is evident that at higher mean ICP, the same bolus of blood will result in higher amplitude of the resulting ICP pulse. The transient increase in intracranial pressure (ΔICP), resulting from the arterial expansion (ΔV) related to the bolus of blood entering the cranium with each heartbeat, can be formulated in the Marmarou equation as:

\[ ICP + \Delta ICP = P_1 e^{k(V + \Delta V)} + P_0 \]

And after reformulation:

\[ \Delta ICP = (e^{k\Delta V} - 1)(ICP - P_0) \]

Assuming that ΔV is constant, there is a linear increase of ΔICP with increasing ICP with the slope \((e^{k\Delta V} - 1)\). This relation has been validated in both humans and animals (20,21). The linear relation holds for baseline and moderately elevated mean ICP. At lower pressures (typically less than 10 mm Hg) ΔICP is constant (21). The relationship of ΔICP and ICP has been termed the "pulsatility curve". And accordingly the pulsatility curve has an ICP independent part at lower ICP and a linear part where ΔICP increase linearly with ICP. The slope, and the ICP below which ΔICP is constant vary between humans (21).

Direct measurement of ICP and ICP pulsatility requires the neurosurgical placement of a sensor inside the skull. The two methods most commonly used in clinical practice are intra-ventricular catheters and intraparenchymal catheter-tip systems (1). After entering the skull, the intracranial pulsations are transferred along the CSF pathway to the subarachnoid space and down through the spinal canal (22,23). This opens up for a potential measurement route via lumbar space, which would be less invasive, and potentially result in fewer complications for patients. Also, assessments would be accessible without a neurosurgical procedure.

Measurements of mean intracranial pressure via lumbar puncture are often performed by neurologists (24). The access of intracranial pulse pressure data via the lumbar route would give the neurologist an additional tool to investigate CSF disorders. This may have special importance in the field of INPH, where abnormal pulsations have been suggested to predict shunt outcome (25). However, this needs to be verified in a randomized clinical trial. Lumbar CSF pressure amplitude measurements may be a useful clinical
tool, but it is important to determine the reliability of the method. In
response to this paper II was conducted.

Idiopathic Normal Pressure Hydrocephalus

Hydrocephalus overview

The first scientific description of hydrocephalus has been ascribed to
Hippocrates (466–377 BC) (26). He also named the disease. The word is
constructed from the Greek words hydor = water, and kefalé = head (26).
There are several proposed classifications of hydrocephalus. The condition
can be classified by the pathological condition leading to CSF accumulation
in the head, either by obstruction of the normal CSF flow pattern, or by
perturbations of CSF production/resorption. Another common classification
is that of communicating/non-communicating hydrocephalus. In the
beginning of the 20th century Dandy and Blackfan injected dye into the
ventricles. If dye was subsequently obtained from a lumbar puncture, the
hydrocephalus was termed "communicating". If no dye was recovered it was
termed obstructive or "non-communicating" hydrocephalus (27). Non-
communicating hydrocephalus can be the result of compression of an
anatomical narrowing (e.g. a tumor), of the ventricular system down to the
point where CSF drains into the subarachnoid space. In communicating
hydrocephalus there is no intraventricular obstruction. Communicating
hydrocephalus can be caused by meningitis, subarachnoid hemorrhage,
traumatic brain injury and brain radiation causing collection of debris or
inflammation in the arachnoid granulations (28,29). To mark that there is a
known cause for the disease, the term "secondary hydrocephalus" is used.
Overproduction of CSF by choroid plexus (hyperplasia/papilloma) leading to
hydrocephalus is very rare (28).

To emphasize that most all hydrocephalus, even "communicating" has an
obstructive component, Rekate in 2008 proposed a definition of
hydrocephalus as "Active distension of the ventricular system of the brain
resulting from inadequate passage of CSF from the point of production
within the cerebral ventricles to the point of absorption into the systemic
circulation." (27).
**Idiopathic Normal Pressure Hydrocephalus**

In 1965, Hakim and Adams described a syndrome of cognitive decline, impaired gait and urinary incontinence in adult patients with communicating hydrocephalus and CSF mean pressure in the normal range (30). The syndrome was later called normal pressure hydrocephalus. To separate this syndrome from hydrocephalus of known cause (secondary hydrocephalus), the term "idiopathic normal pressure hydrocephalus" (INPH) is frequently used.

**Epidemiology**

INPH has a mean age of onset of about 70 years and affects men and women equally (31,32). Since the symptoms of INPH are common in other diseases affecting the elderly, the real prevalence is hard to assess and the condition is likely under diagnosed (32,33). In Norway, Brean et al. conducted a prospective study that sought INPH patients by advertising and asking health workers to refer patients. The prevalence was estimated to 22/100000 and the incidence to 5.5/100 000/year (34). This was not a true population-based study and many INPH patients might not have been referred. Thus, this study likely gives an underestimation of INPH prevalence. Jaraj et al. looked at a representative Swedish population older than 70 years, who had undergone CT-scans and clinical evaluation (32). The estimated prevalence was 0.2 % in the age 70-79 and 5.9% in persons older than 80 years. Israelsson et al. looked at patients who had suffered from a transient ischemic attack and found that 3.9% fulfilled the criteria for INPH (33). These numbers could be compared to the reported incidence of shunt surgery for INPH in Sweden (1996-1998) of 1/100000/year (35). Figure 2 shows the number of implanted shunts per 100 000 inhabitants in adults reported from five out of six neurosurgical centres reporting to the Swedish hydrocephalus register. Approximately 60 % patients in the register have INPH (2011), which gives an incidence for shunt surgery of 2.7/100 000/year. It seems that many INPH patients do not receive treatment, even when conservative estimates of INPH incidence are used.
Pathophysiology of INPH

The pathophysiology of INPH remains unknown. It is however widely agreed that symptoms are the result from a neuronal dysfunction from subcortical structures, for instance caused by a chronic ischemia (31,36). In addition, this is also accompanied by disturbances of the CSF system (31). Here follows a non-exhausting discussion of prevalent pathological findings encountered in the INPH literature.

The resistance to CSF outflow ($R_{out}$) is increased in INPH (37), and have been proposed as the driving force causing ventricle dilation (38). A $R_{out} > 18$ mm Hg/ml/min has been suggested as a limit for shunt surgery (39). Although a feature of the disease, it is not clear why $R_{out}$ is increased, neither how such an increase would cause ventricular dilation (40,41).

Several animal models have highlighted the role of cardiac related intracranial pulsatility as a possible force for ventricular expansion (40). Flow-sensitive MRI have shown increased arterial flow pulsations in INPH.
which has lead to theories of a "water hammer" effect that causes the ventricular dilation (41,43). It has been proposed that age related arterial stiffness and less compliance of the brain leads to larger pressure pulsations in the CSF (44). This is supported by findings that increased blood pressure and blood pressure pulsations is associated with increasing ventricular size (45). Elevated blood pressure is also the largest risk factor for INPH (46). Increased intracranial pulse pressure amplitudes are seen in INPH patients (47), and probing for these is recently proposed as test for selection of shunt candidates (48). Also, the patients improving after shunt surgery have a larger reduction of intracranial amplitudes (49).

**Diagnosing INPH**

Since 2005, formal guidelines exist for diagnosing INPH (50). To reflect variability in diagnostic certainty, the guidelines classify INPH as probable, possible and unlikely. The diagnosis of INPH should be based on history, neurological examination, physiological findings and a neuroradiological investigation. The patient’s response to shunt operation is not considered in the diagnostics, and results of additional “predictive” tests should not be used for diagnostic purposes (51). The clinical presentation of INPH varies greatly from patient to patient, and the symptoms mimic diseases including Parkinson’s, Alzheimer’s, Lewy Body disease, spinal stenosis, and vascular dementia (52,53). In addition, many INPH patients have comorbidities (53,54). Major comorbidities affecting cognition are Alzheimer’s and cerebrovascular disease (54,55). Comorbidities may affect prognosis and outcome after shunt surgery and should be actively sought (53).

**Treatment of INPH**

Many INPH patients improve after insertion of a shunt. The shunt drains CSF, normally from the lateral ventricles, to a low-pressure compartment of the body, normally the abdomen (31). A CSF shunt will decrease $R_{out}$ and has the potential to decrease the ICP as well as the ICP pulse pressure amplitudes (49,56).

Only one randomized controlled trial including a randomized group receiving ligated shunts has been published (57). Patients in this study fulfilled diagnostic criteria of both Binswangers disease and INPH. This study showed neuropsychological and gait improvement in patients with
open shunts (n=7), but not in patients with ligated shunts (n=7). Further, the European multicentre study on INPH showed that as many as 84% of patients (n=142) selected on clinical and radiological criteria alone will benefit from shunt operation (58). Other studies with similar inclusion but other outcome measures have reported lower levels of improvement. Malm et al. found that 72% of patients improved in gait, and Vanneste et al. found that only 29% of patients with communicating hydrocephalus improved after shunt surgery (59,60). Though improved, patients will not regain function in parity with healthy elderly and the extent of potential recovery diminishes with disease duration (61,62). Complication rates in the European study was 28%, with 15% needing additional surgery (58). Complications to surgery include shunt malfunction (20%), subdural hematomas (2-17%), intracerebral hematomas (3%), shunt infections (6%) and seizures (3-11%) (63).

Supplemental tests

To predict which INPH patients would benefit from a shunt operation, several so-called supplemental tests can be used (51). Most common are tests of CSF drainage, where the effect of a shunt operation is simulated. Hakim and Adams first reported clinical improvement after a 15 ml spinal tap, and later Wikkelso et al. reported that 40-50 ml CSF removal was useful for predicting shunt responsiveness (30,64). The specificity of the tap test has been estimated in the range of 33-100% and the sensitivity in the range of 26-100%, largely depending on outcome measures used (51,59,65-68). The standard interpretation is that the tap test has low sensitivity and that patients should not be excluded from shunt surgery based on a negative tap test (51). Sensitivity can be increased by extended lumbar drainage (51), a procedure where typically ten ml CSF per hour is drained over 72 hours. Positive prediction values range between 80-100% (51,69,70). Interestingly, improvement after the test does not seem to be related to change in ventricular volume (71).

The infusion test is a test that give information about a patients CSF system (37). In Umeå infusion tests are performed by an in-house development that has been commercialized (Likvor CELDA®, Likvor AB, Umeå, Sweden). The equipment works by infusing/draining CSF in the lumbar cistern, and simultaneously monitoring the CSF pressure via a second needle in lumbar space. The amount of infused CSF and the pressure response is monitored. From these numbers physiological parameters are calculated (e.g. $R_{out}$). The infusion pattern used in the scope of this thesis is the constant pressure
infusion. The CSF pressure is then regulated and maintained on different constant levels (see figure 3). Average flow to maintain CSF pressure at each level is measured. At entry-level, CSF pressure and associated CSF pressure pulsation amplitudes are measured, without any infusion.

![Graph](image)

**Figure 3.** Graph describing a constant pressure infusion (CPI) protocol where the black lines show the pressure regulation levels. The arrow indicate where the patient is in seated position for CSF sampling. With kind permission from Sara Qvarlander.

Intracranial pulse pressure amplitudes have been suggested to be used for selecting INPH patients for shunt surgery. The technique developed by Eide requires six to 12 hours of continuous ICP measurement. The percentage of time that the amplitude is above certain thresholds is then calculated. Elevated pulse pressure is defined as amplitudes either ≥ 4 mm Hg on average and ≥ 5 mm Hg during more than 10 % of the ICP recording time (72-74).

**Cognitive testing in INPH**

The prevalence of dementia is increasing in the world (75). Proper diagnostics is important and even more so, identifying treatable or reversible causes. Cognitive impairment is one of the cardinal features of INPH and its presence is mandatory for the diagnosis of "probable INPH" according to the INPH guidelines (50). The cognitive disabilities in INPH vary and patients are generally moderately cognitively impaired compared to patients with other dementias (76). The mini mental state examination (MMSE) is a commonly used screening test for global cognitive function (77). The cut off score for dementia is < 24 MMSE points (78). Median MMSE-score for INPH patients at diagnose is 24 out of 30 points and thus 50 % of the patients score above this limit (31,79), indicating the presence of a ceiling
effect when using the MMSE in INPH. This is to some part explained by the MMSE-tests low sensitivity in dementias of subcortical type and makes the test less usable in INPH (76,80). More extensive neuropsychological evaluations show impairments in a majority of INPH patients (81,82).

**Neuropsychological testing in INPH**

The cognitive profile in INPH is characterized as subcortical impairment, with no cortical signs such as apraxia, agnosia aphasia or other focal cortical defects (76). Neuropsychological impairments resemble other subcortical dementias (83-85). Accordingly, neuropsychological studies in INPH have shown impairments in memory (81,86), attention (87,88), psychomotor speed (81,86,89), executive function (81,90,91), and visuospatial skills (92). Impairments in manual dexterity are also seen (81,93,94). Descriptions of mental changes include apathy, emotional indifference and bradyphrenia (95,96). The apathy and bradyphrenia may closely simulate depression (76). Examples of functional impairments are problems with keeping track of appointments, driving, managing finances and taking medications properly (29).

There are a number of neuropsychological studies describing neuropsychological improvement after shunt surgery. The main larger studies will be reviewed in the following.

Malm et al. included 35 INPH patients based on conservative clinical criteria (59). Neuropsychological tests were administered before and three months after shunt surgery. Tests administered were MMSE, Wechsler Adult Intelligence Scale-Revisited (intelligence), the Fuld Object Memory (FOM) Test (consistent retrieval and retrieval failures; memory), the Boston Naming Test (language) and the Figure copy test (visuocognitive skills) from the Alzheimer's Disease Assessment Scale. Significant improvement after shunt surgery was seen in the FOM retrieval failure and Figure copy tests.

Thomas et al. administered neuropsychological tests and MMSE to 42 INPH patients before and at least three months after shunt surgery (86). Neuropsychological tests were the Wechsler Memory Scale (WMS) (verbal memory), Rey Osterrieth Complex Figure (RCFT) (visuocognitive skills and visual memory), Rey Auditory Verbal Learning Test (RAVLT) (memory and learning), Line Tracing test, (psychomotor speed), Trail making test (TMT) B (executive functions), Stroop Color-Word (executive function) and
the MMSE. Different forms of the memory tests were administered at baseline and follow-up to minimize practice effects. Significant improvement in a subtest was defined as one standard deviation (SD) improvement for the patient’s age, sex and educational level. Overall improvement was defined as improvement in more than 50 % of administered tests or a four-point improvement on the MMSE. Overall improvement was seen in 22/42 (52 %) of the patients. Significant improvement was seen in tests of memory and psychomotor speed.

The same definition of overall improvement has been applied in the studies by Duinkerke et al. and Foss et al. (97,98). In the study by Duinkerke 6/10 (60 %) patients improved after shunt surgery. In the study by Foss, the Dementia Rating Scale was administered. Improvement after shunt surgery was seen in 12/27 (44 %) patients. Interestingly patients who improved had significantly higher mean ICP pulse amplitudes in the preoperative evaluation.

Hellström et al. administered the MMSE and neuropsychological tests to 58 INPH patients and 108 healthy individuals (81). The administered tests were Target reaction time (wakefulness/attention), Grooved pegboard test (manual dexterity), tracks (manual dexterity), RAVLT (learning and memory), Swedish Stroop test (executive functions) and Digit span (working memory). INPH patients performed significantly worse than healthy in all administered tasks. Also, patients with vascular risk factors performed worse than those without. There were also far more correlations between test scores for patients, than for healthy controls.

In the follow up study by Hellström, the same tests, with the addition of a simple reaction time test, were administered to 47 patients before and three months after shunt surgery (62). This time results were compared to 159 healthy individuals. Significant improvement was seen in all tests but in the simple reaction time and in Digit span forward test. Though improvement, patients did not reach the performance of the healthy controls in any test. Adopting the same definition as above, overall improvement was seen in 24/47 (51 %) (99). Instead, defining improvement as improved in ≥ four of nine administered tests would yield an improvement rate of 82 % (numbers only presented in dissertation) (99). It was also noted that correlations between tests weakened or disappeared after surgery.

Solana et al performed the largest study on the topic to date. They included 185 INPH patients and administered MMSE and neuropsychological tests before and six months after shunting. Included tests were WMS-revised (information, orientation, visual reproduction and memory), RAVLT
Significant improvement after surgery was seen in all tests except for Digits span forward and TMT B. Significant individual improvement was defined as 1 SD of improvement from the baseline score, adjusted for age, sex and educational level, or by an increase of at least 20 percentile points from the baseline score. Number of improved patients in all tests was around or slightly less than 50 %, except in the TMT B (22 %), Digits span forward (22 %) and Digits span backward tests (33 %). Improvement by ≥ 4 MMSE points was seen in 37/144 (26 %) patients. Most impaired cognitive domains at baseline were seen in tests of memory, executive functions, attention and psychomotor speed. Most marked improvement was seen in tests of memory and psychomotor speed.

Hellström et al. followed up their initial neuropsychological study, with the assessment of a smaller subset of tests (82). The selected tests were the Grooved pegboard test, the Stroop test and RAVLT including delayed recall. The tests were administered to 142 INPH patients at baseline but also 3 months and 1 year after shunt surgery. Scores were compared to 108 healthy controls. Sixty-eight % of patients for the Stroop test and 94 % for the RAVLT could complete the test as baseline. Discriminative ability between INPH and healthy at baseline was demonstrated, with AUC values ranging from 0.86 to 0.95 for included tests. Significant improvement after shunt was seen in all tests. Number of patients who improved more than 1 SD of the distribution of the controls ranged between 14 % in the RAVLT learning test to 62 % in the Grooved pegboard test. The conclusion was that this small set of tests was expedient, diagnostic discriminative and well suited to evaluate changes following shunt treatment (82). Accordingly, these tests, excluding the delayed recall test of RAVLT, were included in a new scale measuring severity and outcome in INPH (100).

One important aspect of neuropsychological testing is that of practice effects. Repeated assessments of the same test typically improve results (101). This would have implications for testing in INPH where repeated testing is common. The typical retest intervals would be 24 hours (tap test), 72 hours (extended lumbar drainage) and 3-6 months (follow up after surgery). Katzen et al. approached the problem of repeated testing by comparing retest results from a healthy control group. In this small study of 12 INPH patients and nine controls, improvement was seen in the TMT B and Symbol digit modalities (102). However, practice effects are affected by population under study (101). Solana et al. specifically studied performance of repeated testing in healthy and INPH patients, and found that practice effects seen in healthy
were not present in INPH patients (103). Also, interestingly Andrén et al. found that a cognitive index from the INPH scale referenced above deteriorated in 33 untreated INPH patients, from 40 points at baseline to 28 points 13.2 (median) months later (61). Implicating that the cognitive decline in INPH is greater than practice effects, at least over this interval. The matter how to evaluate change scores in INPH is to current day unresolved. Adopting the method by Katzen, comparing change scores to those of healthy controls may underestimate effects of treatment as practice effects are larger among healthy. Comparing improvement to baseline scores may in turn overestimate effects, as practice effects may be present. Moreover, complicating the issue further practice effects may depend on baseline performance level, retest interval or cognitive domain (101).

**Computerized neuropsychological testing**

Since the wider adoption of computers there exists computer versions of classical neuropsychological tests (104). A recent review (2014) identified 17 test batteries for testing in the elderly (105). However, paper and pencil tests are still the main cognitive assessment tool (106), which requires a neuropsychologist for selection, administration and interpretation of tests. Computerized batteries may aid in test selection, with pre-selected tests for different diseases, as well as pre-recorded instructions may relieve the neuropsychologist from test administration, making way for more efficient use of resources. Additional advantages for computerized tests may be standardization of instructions, scoring with a high level of precision, cost savings and less influence from examiner bias (105,107). The main criticism has been lacking reports of reliability, equivalence for the response format - particularly for elderly with less computer knowledge, and finally poorly designed computer-person interfaces. Further, redesign of tests may alter validity (107). Computer testing may pose a challenge in the elderly. However, of the different response formats, the touch screen interface has been proposed as the best choice for testing elderly people (108).

There are certainly pros and cons with computerized cognitive testing (105,107), but the idea of having a self-administered cognitive battery holds enough merit to pursue. Especially in INPH where the most often used cognitive screening instrument lacks sensitivity (76,77). Also, standardized self-administered test may facilitate research comparability and collaboration. In the response to this, in addition to the other potential advantages discussed above, a computerized neuropsychological test battery customized for INPH was designed and evaluated in papers III and IV.
**Test development**

Before deploying a new test battery several factors must be examined. It must be assured that tests are valid, reliable and that the intended patient group can complete the tests.

*Validity* can be defined as "the degree to which an instrument truly measures the construct it purports to measure" (109). There are a plethora of concepts relating to different aspects of validity: Face validity, the degree to which a construct looks as an adequate reflection of the construct to be measured; Criterion validity, do the scores reflect a gold standard? Ecological validity, the degree to which scores in the test says something about the behaviour in a more natural environment (110). Another aspect of validity is responsiveness; how well an instrument reflect changes in the underlying construct (109). It has been argued, because of the different response format, that discriminant validity, the degree to which different tests do not co-vary, is important in computerized testing (105).

The *reliability* of an instrument is defined as "the degree to which a measurement is free from measurement error" (109). If the measurement error is too large, statements about impairments or improvements of cognitive functions become unreliable. The concept is also related to validity in that covariation of scores from different measurement methods is limited by the degree of error in the measurements.

Neuropsychological assessment provides non-invasive means of assisting in diagnostics, tracking disease progression and assessing efficacy of treatment in INPH. A self-administered neuropsychological test battery with preselected tests may increase standardization and accessibility for cognitive testing. A prerequisite is that implemented tests are reliable, valid, and responsive. Also, to be useful they should show good completion rates.

**Research rationale**

There are an increasing number of studies about many aspects of INPH, see figure 4. Each study represents a tremendous amount of effort. However, what can be concluded from these studies is often limited by small study
samples. Pooling over several studies or comparing results is difficult, as different measurements methods are often used.

Measurements often provide the basis for diagnosis and medical interventions. There may be limitations due to low quality of measurements and the use of measurements that are insufficiently characterized. It is therefore important that measurements are thoroughly evaluated and their limitations reported and understood.

In response to this, the studies in this thesis contribute by exploring limitations in ICP measurements. Specifically the PI method of estimating ICP and the pulse pressure amplitude measurement via the lumbar route are explored. Also, aiding standardization of cognitive measurements, a self-administrating computerized neuropsychological test battery is developed. The hypothesis is that a standardized computerized test battery would be useful in the cognitive assessments of INPH-patients.

![Figure 4. Pubmed search of "Idiopathic Normal Pressure Hydrocephalus". The graph shows number of publications per year.](image-url)
AIMS

I. To validate the method of estimating intracranial pressure via transcranial Doppler sonography pulsatility index of middle cerebral artery blood flow waves.

II. To investigate whether cardiac related intracranial pulse pressure amplitudes in the brain correspond to those measured in lumbar space.

III. To develop and evaluate validity and reliability of a computerized neuropsychological test battery customized for INPH.

IV. To assess the completion rate for computerized tests in INPH patients and to test the ability to detect cognitive impairment at baseline and improvement after a shunt surgery in INPH patients.
MATERIALS AND METHODS

Ethical approval

The Regional Ethical Review Board in Umeå approved all the studies presented in this thesis.

Patients and controls

Table 2. Patient cohorts used in this thesis.

<table>
<thead>
<tr>
<th>Paper</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>INPH n=10</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Healthy elderly n=44</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Cognitively impaired n=28</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Hydrocephalus n=40</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>INPH(^a) n=26</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>INPH shunt surgery n=31</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Of the 40 patients in the Hydrocephalus group in paper III, 26 were diagnosed with INPH.

INPH patients of paper I and II

Ten INPH patients (two women) were included, mean age 72.4 years (range 55-78). MRI revealed communicating hydrocephalus, with dilated ventricles (all had Evens Index >0.3), without cortical atrophy or severe white matter lesions. Diagnostics included clinical characteristics, routine laboratory tests, MMSE (77), and video taped gait tests.
Overnight intra-parenchymal pressure monitoring was performed. A lumbar CSF infusion test and tap test was performed in the morning the day after implantation of the catheter. While the infusion test was running, TCD of the MCA and lumbar pressure monitoring was performed. All patients subsequently received shunts. The studies were part of a larger project with several previous publications (24,36,111).

**Cognitively impaired patients of paper III**

A research nurse screened the neurological ward for patients. Inclusion criterion was a MMSE between 20 and 30. The only exclusion criterion was impaired motor function (e.g. palsy). Thirty patients agreed to participate and two patients were excluded because of inability to perform either conventional or computerized testing, resulting in twenty-eight included patients. See table 3 for demographics.

**Hydrocephalus patients of paper III**

Forty consecutive patients referred to our department because of suspected INPH were included. All patients had MRI-verified communicating hydrocephalus, see table 3 for demographics. A review of the charts revealed that 26 patients fulfilled the criteria for INPH according to American and European guidelines (50).

**Healthy elderly of papers III and IV**

Forty-five healthy elderly were recruited via an ad in the local newspaper in Umeå. Inclusion criteria were an age between 60 and 82 years old. The participants were confirmed healthy regarding medical history and clinical examination including on-going medication, physical and neurological examinations, electrocardiography, blood pressure, body mass index, MMSE and MRI. Exclusion criteria were: MMSE<28, disease of the nervous system, medication affecting the nervous system (such as benzodiazepine or antidepressants), anticoagulants, ischemic heart disease and diabetes. More than one of: hypertension, smoking or hyperlipidemia, was also an indication for exclusion. One participant was excluded due to technical
problems with the touch screen, resulting in 44 included subjects. Demographics are displayed in table 3.

**Table 3. Characteristics of the study populations of paper III**

<table>
<thead>
<tr>
<th></th>
<th>Healthy elderly</th>
<th>Cognitively impaired patients</th>
<th>Hydrocephalus patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y Median (range)</td>
<td>69 (60-79)</td>
<td>71 (56-86)</td>
<td>72 (50-85)</td>
</tr>
<tr>
<td>Numbers (n)</td>
<td>44*</td>
<td>28</td>
<td>40</td>
</tr>
<tr>
<td>Male/Female, %</td>
<td>41/59</td>
<td>50/50</td>
<td>63/37</td>
</tr>
<tr>
<td>Education y, Median (range)</td>
<td>11.5 (6-22)</td>
<td>10 (6-15.5)</td>
<td>8 (6-20)</td>
</tr>
<tr>
<td>Computer knowledge %</td>
<td>60</td>
<td>50</td>
<td>53</td>
</tr>
<tr>
<td>Color blind %</td>
<td>0</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>MMSE, Median (range)</td>
<td>&gt;28</td>
<td>26 (20-30)</td>
<td>26 (18-30)</td>
</tr>
<tr>
<td>GDS, Median (range)</td>
<td>0 (0-6)</td>
<td>3 (0-10)</td>
<td>4 (0-19)</td>
</tr>
</tbody>
</table>

MMSE = Mini Mental State Exam, GDS = Geriatric Depression Scale, *Two tests (Four-finger tapping and Ten word list learning) were redesigned during the study and only 26 of the 44 participants of the healthy elderly took the slightly modified battery. †The subjects were asked, "Do you have computer knowledge, yes or no".

**INPH patients of paper IV**

Thirty-two INPH patients were recruited from three centres (Umeå, Linköping and Aalborg). Patients were included if older than 60 years, having INPH according to international guidelines (50), and were planned for surgery. Exclusion criteria were secondary NPH or medical condition preventing neuropsychological testing (e.g. blindness, deafness). Demographics of included patients are displayed in table 4. One patient was excluded due to non-functioning shunt, resulting in 31 included patients.
There was an overlap in recruitment, and four patients were included in both papers III and IV.

**Table 4.** Characteristics of INPH patients in paper IV.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INPH</strong></td>
<td><strong>n=31</strong></td>
</tr>
<tr>
<td>Age, y Median (range)</td>
<td>74 (64-85)</td>
</tr>
<tr>
<td>Male/Female, %</td>
<td>84/16</td>
</tr>
<tr>
<td>Education y, Median (range)</td>
<td>8 (6-17)</td>
</tr>
<tr>
<td>Computer knowledge % Yes</td>
<td>35</td>
</tr>
<tr>
<td>Color blind %</td>
<td>6</td>
</tr>
<tr>
<td>MMSE, Median (range)</td>
<td>26 (19-30)</td>
</tr>
<tr>
<td>GDS, Median (range)</td>
<td>4 (0-13)</td>
</tr>
</tbody>
</table>

MMSE = Mini Mental State Exam, GDS = Geriatric Depression Scale.

**Simultaneous infusion test, transcranial Doppler and pressure measurements**

In papers I and II, the patients performed simultaneous measurements of MCA blood flow velocity, blood pressure, ICP and lumbar CSF pressure, while ICP was regulated by lumbar infusion/withdrawal of CSF. An intraparenchymal catheter tip sensor (Codman MicroSensor™ Johnson & Johnson Professional, Raynham, MA, USA) was inserted into brain parenchyma close to the frontal horn of the right ventricle, and an overnight registration of ICP was performed. On day two, with the intracranial pressure sensor still in place, two needles were inserted in the lumbar subarachnoid space. One needle was used for pressure monitoring via a fluid catheter system (Becton Dickinson, Franklin Lakes, NJ, USA) and the other for CSF volume alternation. The CSF pressure was measured at the midpoint.
between the highest and lowest points of the head, with the patient in the supine position.

For CSF volume alternation, an in house developed infusion apparatus was used. A constant pressure method was used, meaning that the infusion was regulated to keep mean ICP stable at pre-set levels. First a baseline recording was performed, and then intracranial pressure was increased in two or threes steps up to 45 mm Hg, before pressure was released back to baseline again. The recording was ended with CSF withdrawal to obtain an ICP of zero mm Hg, see figure 5. On each pressure level a 30-90 seconds TCD recording from the left middle cerebral artery was performed.

![Diagram](image)

**Figure 5.** Constant pressure infusion in one patient. The pressure was recorded at baseline. Then pressure was increased in three steps (1st - 3rd excess), and finally released to baseline again (Relax). The recording ended with CSF withdrawal to zero ICP (Drain). The magnification shows intracranial and CSF pressure waves.

All data was exported and analyzed offline in Matlab® (Mathworks Inc., Natick, MA, USA). A program for automatic waveform analysis was developed. The program identified diastolic minima and systolic maxima for blood pressure, TCD velocity, as well as intracranial and lumbar CSF pressure waves, see figure 6. After pulse wave identification, the amplitude of each pulse was calculated as systolic maxima minus previous diastolic

---

The delay between intracranial and CSF pulses was calculated as time difference between diastolic minima. Also the fast Fourier transform (FFT) was applied to the data. The amplitude of the fundamental frequency, and the two successive over-tones was calculated for consecutive 6-second time windows.

![Blood pressure and Intracranial pressure](image)

**Figure 6.** In the figure, waveforms from intracranial pressure and blood pressure are displayed. The program calculated the fundamental frequency for blood pressure data and looked for successive local minima within this interval. Corresponding minima for ICP was sought close these points. Systolic maximum was calculated as the maximum value between two minima. Asterisks indicate the limits of the pulse wave found by the program.

### Mathematical model

In paper I a mathematical model was invoked (112-114). The model is a complex physiological model describing interaction of a number of mechanisms that each are described by a differential equation. The model incorporates the CSF circulation, intracranial pressure volume relationship, and cerebral hemodynamics, including auto regulatory mechanisms. The model has previously been invoked to study TCD waveforms (115).

Input to the model was a blood pressure waveform from one of the patients. The differential equations were solved with numerical methods in Matlab. Results were also verified in a Simulink model. In the model, the in-vivo experiment was reproduced, simulating a constant pressure infusion investigation, see figure 7. The simulations were done with different values for physiological parameters and the resulting waveform of the MCA was
analyzed with the program described above. Diastolic minima and systolic maxima were identified and the pulsatility index for each pressure level was calculated.

![Graph showing intracranial pressure over time and infusion rate over time.](image)

**Figure 7.** Simulated infusion test with the mathematical model. Above is the resulting intracranial pressure and below the infusion rate as function of time.

**Bench test**

In paper II a bench test comparing the Codman catheter tip sensor with the Becton fluid catheter system was performed. The systems were set up in the same way as the in-vivo measurements, but connected to a pressure waveform generator (Model 601A Blood pressure system calibrator, Biotech Instruments, Inc. Burlington, VT, USA) for simultaneous measurements. A pressure waveform typical for the patients was modified and fed to the systems with different amplitudes, heart rates and mean pressures.
Neuropsychological test battery

In papers III and IV a computerized neuropsychological test battery was developed and evaluated. A Pubmed literature search was performed, finding the most frequently used neuropsychological tests that were also sensitive to the cognitive profile of INPH patients. The tests were selected for brevity, feasibility for implementation on a touch screen and to cover the cognitive domains affected by INPH. Also, tests had to be in the public domain, or the implemented version had to be sufficiently different not to infringe any copyright. This was in effect solved by the computer adaptation itself.

Software design

Tests were implemented using the computer languages JAVA (116) and Adobe Flash (117). The Flash language was used for making the user interface, with graphics, buttons and animations with pre-recorded sound presenting the tests. JAVA was used to implement storage of test results and sound from the microphone to the hard drive and the generation of a test report. A pilot test battery was developed and tested on personnel at our department. After this, bugs were corrected and the tests that did not work well in practice were redesigned (e.g. Four-finger tapping on the touch screen was replaced by a numeric keyboard).

Computerized General Neuropsychological INPH Test (CoGNIT)

The CoGNIT battery includes tests of memory, attention, psychomotor speed, executive functions, and manual dexterity. Here follows a description of implemented tests. Screen shots from some tests are displayed in figure 8.

1. Two choice reaction test (attention). The subject is instructed to keep the stylus over a cross in the middle of the screen, and then press one of two buttons indicated by an arrow (see figure 8A). The arrow appears after a random interval of 5 to 15 seconds. Median reaction time over 20 trials is used as the test score.
2. **Stroop congruent colors (psychomotor speed)** (118). Two buttons of different colors (red, green, yellow or blue) are displayed. One corresponds to the name of the color presented in black text (see figure 8B). The subject is asked to press the button with the color corresponding to the text. A new text and buttons then appears after a 2 seconds delay. Median reaction time of 50 words is measured.

3. **Stroop incongruent colors (executive functions)** (118). Two buttons of different colors (red, green, yellow or blue) and colored text are displayed (see figure 8C). The subject is asked to press the button with color corresponding to the color (not the text) of the text. A new text and buttons then appears after a 2 seconds delay. If instructions are not followed while the test is running, the observer has instructions to verbally correct the patient at a maximum of two times. Median reaction time of 50 words is measured. In **paper III**, if the patient made more than 50% errors, the test was regarded as failed. This criterion was loosened for **paper IV**, where 10 correct answers were sufficient to complete the test.

4. **Trail making test A (psychomotor speed)** (119). The subject is asked to press numbers in consecutive order (1-25) displayed on the screen. Time to completion is measured.

5. **Trail making test B (executive functions)** (119). The screen displays letters (A-L) and numbers (1-13) (see figure 8D). The subject is asked to press buttons by alternating between letters and digits (1-A-2-B-3-C...). Time to completion is measured.

6. **Ten-word-list (memory)** (120). The subject is asked to remember 10 consecutive words presented with text and pre-recorded announcer speech. The words are randomly drawn from a pool of the 50 most common Swedish nouns. The subject is then asked to repeat the wordlist into a microphone (see figure 8E). The task is repeated three times. The score is the sum of correctly remembered words, summed over all three trials.

7. **Delayed recall (memory)** (120). After approximately 20 minutes of distracter tasks, the subject is asked to repeat the words from the list-learning task. The number of correctly recalled words is used as the score.

8. **Delayed recognition (memory)** (120). The subject is asked to recognize the 10 words from the list-learning task among 10 distracter words. Words are presented consecutively. The test score is calculated as the number of correct responses minus errors.
9. **Figure copy task (visuospatial skill)** (120). The subject is asked to copy a cube displayed on the screen. The drawing is manually scored as correct or incorrect. The figure is regarded as correct if the size is correct and all lines are present.

![Figure 8](image)

**Figure 8.** Screen shots from some of the included tests. A, Two-choice reaction test, B Stroop congruent colors test, C Stroop incongruent colors test, D, Trail making test B, E Ten-word list test. F Four-finger tapping test.
10. **Four-finger tapping (manual dexterity)** (121). The subject is asked to tap on a small numeric keyboard with fingers 2-5 of the dominant hand (see figure 8F). The correct order of tapping is (digits): 2-3-4-5-4-3-2-3-4-5-4-3.... The number of correct taps during 10 seconds is measured. The task is repeated for 5 times and the sum of taps over all trials is the score.

11. **Geriatric Depression Scale (depression)** (122). The Geriatric depression scale is a 20 questions instrument intended to measure symptoms of depression in the elderly. The questions are presented on the screen in consecutive order and the subject is required to give answers by pressing buttons marked yes or no.

**Statistics**

Significance threshold for all statistical tests was set to 0.05. Statistical softwares used were R version 2.13 in **paper I**, and SPSS (SPSS Inc, Chicago, IL, USA) in **papers II-IV** (version 18 in paper II and version 21 in papers III-IV).

In **paper I** correlations and linear regressions between ICP and PI was performed both on lumped data and on data stratified by individual patients. Confidence and prediction intervals were calculated.

In **paper II** the relation between intracranial and lumbar pressure amplitudes, as well as bench test data from the two sensors, were explored by Bland Altman plots (123). Spearman correlations were consistently used. Differences between intracranial and lumbar amplitudes (ΔICP - ΔLP) were explored in a general linear model with mean intracranial pressure (k*ICPmean) as covariate and a patient dependent factor (mpat). The model was invoked to explain amplitude difference in time domain amplitudes (ktime), amplitudes calculated from frequency domain fundamental frequency and two over tones (kfund, kOT1 and kOT2). Also the same model was used to explain the delay between intracranial and lumbar pressure waves (kdelay).

In **paper III** reliability was explored by Pearson correlation (r) between test and retest scores of healthy elderly. The standard error of measurement SEm was calculated for applicable tests, as $SEm=SD*sqrt(1-r)$, where SD is the standard deviation of the test scores. The SEm gives an error band around a single given score, and the true score is approximately within $±2\ SEm$ with 95% confidence. Practice effects between test and retest were explored with
paired T-tests when the normality assumption was met otherwise the Wilcoxon signed rank test was used. Association between corresponding scores from computerized and conventional, and between different computerized tests was explored by Spearman correlations.

In paper IV improvement after shunt operation was explored by the Wilcoxon signed rank test. The individual change (improved or worse) in a test was defined as change in ability to complete the test, or scores changes exceeding "SEm of the score difference" (SEm_{diff}), which was calculated from values in paper III, as:

\[ SEm_{diff} = \sqrt{2} \times SEm \]

Overall cognitive level was explored by summing Z-scores of all tests. Standard deviations and means were obtained from INPH patients providing data to each test. In tests where patients failed, the Z-score was set to the lowest value of those patients who managed to complete the test.
RESULTS

Transcranial Doppler pulsatility index and intracranial pressure

The primary aim of paper I was to evaluate if TCD is a feasible method to assess ICP. The overall regression line relating pulsatility index (PI) to ICP was ICP=23*PI+14 (R²=0.22, p<0.01). The 95% prediction interval for ICP given a PI measurement, was in the order of ±25 mm Hg, see Figure 9. Linear regressions for data stratified by patient showed great intra-individual variability. For instance patient 3 (ICP=82*PI-32, R²=0.86) and patient 4 (ICP=48*PI+3, R²=0.85) had a factor 2 slope difference.

Figure 9. The regression line between PI and ICP. The 95% confidence and prediction intervals are displayed.

Simulating the mathematical model showed that, with no changes in physiological parameters, PI truly reflected ICP. However, altered physiological variables such as mean arterial pressure, arterial pressure amplitude, autoregulation and compliance of the middle cerebral artery distorted the relationship, see Figure 10.

Figure 10. Simulations of the ICP-PI relations’ sensitivity to alternations in physiological parameters. Altered parameters are (A) mean arterial pressure (MAP), (B) arterial pulse pressure amplitude, (C) autoregulatory gain, and (D) middle cerebral artery (MCA) compliance ($K_{MCA}$). An autoregulatory gain parameter equal to 1 means normal autoregulation, whereas zero implies non-functioning autoregulation. A low $K_{MCA}$ value indicates a more compliant vessel.

Intracranial and lumbar pressure waves

The aim of paper II was to determine if ICP pressure wave amplitudes were accurately assessed via lumbar puncture. To account for hardware differences in intracranial and lumbar measurement systems, also a bench test was performed.

---

**Baseline**

At baseline pressure (mean ICP = 18.7 mm Hg, mean ICP amplitude = 6.1 mm Hg) the intracranial amplitudes preceded and exceeded amplitudes measured the lumbar route (ΔAmp = 0.9 mm Hg, SD=1.0, p<0.05; Delay=39 ms). Bench test data showed small measurement related differences at amplitudes corresponding to baseline readings (ΔAmp\textsubscript{sensors} < 0.2 mm Hg).

**Elevated ICP**

The correlation between intracranial and lumbar pressure amplitudes was high (r=0.98, p<0.001), see figure 10. At higher pressures the lumbar readings increased more, and even exceeded intracranial amplitudes at the highest pressure level (mean ICP = 44.3 mm Hg, mean ICP amplitude = 21.7 mm Hg, ΔAmp = -0.2). The lumbar delay were still present but smaller (Delay = 18 ms). These results were supported by the GLMs (k\textsubscript{time} = -0.033, P < 0.01; k\textsubscript{delay}=-0.63, P < 0.001). The GLMs also revealed variability in the patient dependent factor m\textsubscript{pat}.

![Figure 11](image)

**Figure 11.** In (A) the agreement between ΔICP and ΔLP is displayed. There was an excellent correlation (r=0.98, r<0.001). (B) shows the amplitude difference as function of mean amplitude (r=-0.47, p<0.01). Numbers represent patients.

Bench test data revealed a relative overestimation by the lumbar sensor at the higher-pressure amplitudes (Figure 11 A and B). Higher heart rate gave

---

larger sensor discrepancy (Figure 11 B). Differences in mean pressure did not affect sensor readings.

**Figure 12.** Bench test amplitude difference as function of mean pressure amplitude calculated from time and frequency domain data. Data represents means for six measurements with different sensors. Figure (A) displays results for a heart rate of 60 bpm ($r = -0.73, p < 0.001$ and $r = 0.73, p < 0.001$) and (B) 120 bpm ($r = -0.94, p < 0.001$ and $r = 0.13, p = 0.11$).

**Time versus frequency analysis**

In the in vivo experiment, amplitudes calculated from the fundamental frequency, consistently underestimated time domain analysis. Looking at ICP amplitudes, the difference was 2 mm Hg at baseline and 3.8 mm Hg at the highest-pressure level. A frequency analysis revealed that not the fundamental, but the first two overtones increased with higher mean pressures in lumbar compared to intracranial readings ($k_{fund} = -0.002, P = 0.85$; $k_{OT1} = -0.017, P < 0.01$ and $k_{OT2} = -0.015, P < 0.01$). Bench test data also revealed that amplitudes calculated from the fundamental frequency were less affected by sensor differences.

---

CoGNIT

The aim of papers III and IV was to implement commonly used neuropsychological tests for INPH into a computerized format, and to evaluate the reliability, validity, responsiveness to shunt surgery and completion rate for INPH patients. Also, scores for healthy elderly were collected for comparison.

Validity

Validity was explored by correlations between corresponding conventional and computerized test scores in patients with cognitive impairment. Results are displayed in table 5. There were significant correlations between all computerized and corresponding conventional tests ($r=0.49$-$0.87$).

**Table 5.** Correlation between conventional and computerized tests.

<table>
<thead>
<tr>
<th>Test</th>
<th>Correlation computer-</th>
<th>Significance of correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroop congruent</td>
<td>0.82</td>
<td>$&lt;$0.001</td>
</tr>
<tr>
<td>Stroop incongruent</td>
<td>0.76</td>
<td>$&lt;$0.001</td>
</tr>
<tr>
<td>Ten word list</td>
<td>0.66</td>
<td>$&lt;$0.001</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>0.72</td>
<td>$&lt;$0.001</td>
</tr>
<tr>
<td>Delayed recognition</td>
<td>0.49</td>
<td>$&lt;$0.01</td>
</tr>
<tr>
<td>Trail making test A</td>
<td>0.85</td>
<td>$&lt;$0.001</td>
</tr>
<tr>
<td>Trail making test B</td>
<td>0.83</td>
<td>$&lt;$0.001</td>
</tr>
<tr>
<td>Figure copy task</td>
<td>0.54</td>
<td>$&lt;$0.01</td>
</tr>
</tbody>
</table>

*No corresponding paper and pencil test exist for the Two-choice reaction and Four finger tapping tests.*
Divergent validity was explored by correlation between different computerized tests. Results are displayed in table 6. Significant correlations exist between tests assessing the same domain. Also, tests with a strong motor component correlate e.g. the Stroop tests, the Trail making tests and the Four-finger tapping test. A correlation was also seen between the Delayed recognition and Figure copy tests.

Table 6. Correlation matrix of healthy elderly performance at first computer test session.

<table>
<thead>
<tr>
<th>TESTS</th>
<th>Two-choice reaction</th>
<th>Stroop congruent</th>
<th>Stroop incongruent</th>
<th>Trail making A</th>
<th>Trail making B</th>
<th>Ten-word list</th>
<th>Delayed recall</th>
<th>Delayed recognition</th>
<th>Four-finger tapping</th>
<th>Figure copy test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-choice reaction</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroop congruent</td>
<td>0.63</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroop incongruent</td>
<td>0.4</td>
<td>0.43</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trail making test A</td>
<td>NS</td>
<td>0.45</td>
<td>0.46</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trail making test B</td>
<td>NS</td>
<td>NS</td>
<td>0.48</td>
<td>0.73</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ten-word list</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed recall</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>0.61</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Delayed recognition</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>1</td>
</tr>
<tr>
<td>Four-finger tapping</td>
<td>NS</td>
<td>-0.49</td>
<td>-0.56</td>
<td>-0.39</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>-0.44</td>
<td>NS</td>
<td>1</td>
</tr>
<tr>
<td>Figure copy test</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>0.44</td>
</tr>
</tbody>
</table>

NS = Non significant.

**Reliability**

Reliability was explored by correlations between repeated computerized testing in healthy subjects. Test-retest correlations and tests of improvement between test and retest are displayed in table 7. The mean test-retest interval was 21 days and correlations between test and retest were >0.7 in 8 out of 10 tests. The Figure copy test had a correlation of 0.57 and the Ten-word list test had a correlation of 0.67. There were significant improvements at retesting in 5 tests.
Table 7. Results from the reliability study.

<table>
<thead>
<tr>
<th>Computer test</th>
<th>Improvement test-retest</th>
<th>Correlation test-retest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two choice reaction</td>
<td>NS</td>
<td>0.75</td>
</tr>
<tr>
<td>Stroop congruent</td>
<td>NS</td>
<td>0.74</td>
</tr>
<tr>
<td>Stroop incongruent</td>
<td>&lt;0.01</td>
<td>0.83</td>
</tr>
<tr>
<td>Ten word list</td>
<td>&lt;0.001</td>
<td>0.67</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>NS</td>
<td>0.74</td>
</tr>
<tr>
<td>Delayed recognition</td>
<td>NS</td>
<td>0.70</td>
</tr>
<tr>
<td>Trail making test A</td>
<td>&lt; 0.05</td>
<td>0.87</td>
</tr>
<tr>
<td>Trail making test B</td>
<td>&lt; 0.05</td>
<td>0.83</td>
</tr>
<tr>
<td>Figure copy task</td>
<td>NS</td>
<td>0.57</td>
</tr>
<tr>
<td>Four finger tapping</td>
<td>&lt;0.001</td>
<td>0.90</td>
</tr>
</tbody>
</table>

NS = Non significant

Ability to complete the test

In paper III 78% of consecutively evaluated hydrocephalus patients completed the battery with one or zero failed tests. The corresponding number from paper IV during the preoperative investigation of INPH patients was 81%. The most frequently non-completed tests were Trail making test B, Stroop incongruent colors test and the Four-finger tapping test.
**Performance of untreated INPH patients**

INPH patients recruited in papers III and IV performed significantly worse compared to healthy on all administered tests. Interquartile ranges between INPH patients and healthy were non-overlapping in all tests but a small overlap in the Trail making test B in paper III (see figure 13). The overall profile of cognitive impairment was similar between the two INPH cohorts. In paper IV it was shown that vascular risk factors affected preoperative scores negatively.

![Figure 13](image.png)

**Figure 13.** Performance of INPH patients recruited in papers III and IV. Results are normed so that 100% represents median scores of healthy controls. Bars indicate interquartile ranges.

**Neuropsychological improvement after shunt**

The aim of paper IV was to evaluate the effects of shunt surgery on CoGNIT performance in INPH patients. Mean delay from cognitive testing to shunt
surgery was 5.6 months. Mean follow up was 4.2 months after surgery. Improvement in tests on group and individual level are displayed in table 8.

After shunt surgery, patients were still impaired compared to healthy with the exception of the Four finger-tapping test. Significant improvement was seen in all cognitive domains: Memory (Ten-word list); Attention and psychomotor speed (Two choice reaction test, Stroop congruent colors); Executive functions (Stroop incongruent colors) and Manual dexterity (Four-finger tapping). Improvement at trend level was also seen in the Delayed recall test (p=0.16). Improvement in summed Z-scores was seen in 14 patients (45%).

Table 8. Neuropsychological improvement after shunt surgery.

<table>
<thead>
<tr>
<th>Test</th>
<th>Improvement</th>
<th>Individual changea</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P-value</td>
<td>I/U/W</td>
</tr>
<tr>
<td>Two choice reaction</td>
<td>&lt;0.01</td>
<td>16/11/4</td>
</tr>
<tr>
<td>Stroop congruent</td>
<td>&lt;0.05</td>
<td>14/10/7</td>
</tr>
<tr>
<td>Stroop incongruent (n=19)</td>
<td>&lt;0.05</td>
<td>14/11/6</td>
</tr>
<tr>
<td>Ten word list</td>
<td>&lt;0.01</td>
<td>13/14/4</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>NS</td>
<td>11/13/7</td>
</tr>
<tr>
<td>Delayed recognition</td>
<td>NS</td>
<td>9/18/4</td>
</tr>
<tr>
<td>Trail making test A (n=30)</td>
<td>NS</td>
<td>11/9/11</td>
</tr>
<tr>
<td>Trail making test B (n=22)</td>
<td>NS</td>
<td>10/11/10</td>
</tr>
<tr>
<td>Four finger tapping (n=12)</td>
<td>&lt;0.01</td>
<td>11/13/7</td>
</tr>
<tr>
<td>Summed Z-scores</td>
<td>&lt;0.01</td>
<td>14/12/5</td>
</tr>
</tbody>
</table>

I = numbers improved, U = Unchanged, W = Worse, NS = Non significant

I = numbers improved, U = Unchanged, W = Worse, NS = Non significant
DISCUSSION

**Paper I** validated the non-invasive transcranial Doppler pulsatility index (PI) method of estimating ICP. The experiment showed that PI was an unreliable predictor of ICP. In a mathematical model, it was shown that PI was affected by various physiological parameters apart from ICP.

In **paper II**, pressure wave amplitudes were simultaneously measured in the brain parenchyma and via a lumbar puncture in INPH patients. Main finding was that at resting pressure, amplitudes were slightly attenuated in lumbar CSF compared to in the brain parenchyma. In summary, lumbar pressure amplitude measurement with a fluid catheter system is an alternative to intraparenchymal ICP measurement, but the thresholds for what should be interpreted as elevated amplitudes need to be adjusted.

In **papers III and IV** a computerized neuropsychological test battery for INPH was developed and evaluated (COmputerized General Neuropsychological INPH Test = CoGNIT). Tests were chosen after a literature review and correlated significantly with corresponding paper-and-pen versions. In **paper III** it was shown that reliability was good, and the completion rates in INPH satisfying. In **paper IV** it was established that CoGNIT showed impairment in untreated INPH patients compared to healthy controls and improvement after shunt surgery was detected in all assessed cognitive domains. Consequently CoGNIT is a new instrument in the INPH field, for research but also for daily patient care.

**Pulsatility index and ICP**

The results of **paper I** refutes the pulsatility index method of predicting ICP. The prediction interval for ICP given a PI reading was in the order of ±25 mm Hg, and henceforth too wide to be of clinical utility. The large prediction interval of lumped data was explained by a large variability in the PI/ICP-relation between patients. The mathematical simulations indicated that differences may be due to physiological variation. Prior to publication of **paper I**, there were many positive publications regarding the potential usefulness of the pulsatility method based largely on data from patients with traumatic brain injury (14-17), even though regression lines for the ICP/PI-relation differed largely between studies (15,17).
A limitation with the TCD technique is that it is affected by operator dependent factors, which may explain differences between studies (124). Also differences in constitution of the temporal bone window make TCD data inaccessible in approximately 10-15 % of patients(125).

The negative result from our study may be due to the clinical material, where hydrocephalus patients may show a less perfect PI-ICP correlation. However, after our publication, one large study with 290 TBI patients with continuous ICP monitoring also looked at prediction intervals of ICP from TCD PI data. The prediction interval was between ±15-35 mm Hg, depending magnitude of PI (126). Also, our results from the mathematical simulations has been confirmed showing that many other variables than ICP affect PI, e.g. arterial pulse pressure, heart rate and cerebral perfusion pressure (127).

The strength of paper I is that PI was measured over a wide range of ICPs in every patient. From this data it could be concluded that the wide prediction interval of lumped data stem from differences between individuals. The differences were so large that it is not possible to transfer a general PI-ICP relation to the individual level.

One obvious limitation with paper I is the low number of participants. However, even including more patients, the variation from the first patients would remain, but of course give better estimations of the prediction interval of ICPs.

A tool for non-invasive ICP measurement would be very useful in clinical practice. However, with such a complex system, with a multitude of interactions such as the cerebrovascular hemodynamic system of the brain, it would be surprising if a simple index such as PI from the middle cerebral artery would predict ICP with high accuracy. Extraordinary claims require extraordinary evidence and even more so when the utility is with critically ill patients. The more enthusiastic reports on this topic should be interpreted carefully, and future studies should include prediction intervals for derived ICP.

**Pulse pressure amplitudes of the ICP**

The pulsatile stress from each heartbeat will eventually exhaust predominantly the non-regenerating materials of the body. Microfractures and degeneration in the walls of aorta is seen in normal ageing, leading to dilation and stiffening (128). There is then a breakdown of the arterial
Windkessel that leads to larger flow pulsations, transmitted predominantly to large flow organs such as the brain and the kidney (128). Also, the compliance of the craniospinal compartment decreases with age (129,130). An increased pulsatility transmitted from the larger arteries lead to microvascular damage with smaller arterial lumen and less mean flow (131). Larger arterial pulsatile flow, together with decreased compliance of the brain will result in larger ICP pulsatility that may indirectly affect the brain (132). Recent studies in healthy elderly has linked larger arterial and intracranial pulsations to decreased cognitive performance, smaller total and regional brain volumes, larger ventricles and more white matter lucencies on MRI (132-134). INPH patients have larger ICP amplitudes and the assessment of those is proposed as a predictive test for shunt surgery (72,74). ICP amplitudes would be far more accessible if measurable the lumbar route.

In paper II, there was a good correlation between intracranial and lumbar pressure amplitudes. Relying on lumbar measurement will underestimate true ICP amplitudes by approximately 0.9±1 mm Hg at baseline ICP. This underestimation holds for the clinically most relevant range, and is probably due to dampening of the pulse waves traveling down the spinal canal. The standard deviation of amplitude differences may to a smaller degree be due to measurement error, and are most likely due to individual differences.

These limitations is important to take note of e.g. using lumbar measurements in the test for selecting INPH-shunt candidates proposed by Eide (74). Maybe, recording of resting lumbar CSF pressure amplitudes can only reliably be used in a prognostic test if they are very large. Accordingly Eide et al. found that lumbar CSF pressure amplitude at baseline was <2 mm Hg in 17/42 of patients who had mean ICP amplitude ≥4 mm Hg in overnight intraparenchymal registrations (72). Interestingly, Eide also found that prediction of patients with large ICP amplitudes in an overnight registration was increased if lumbar pulse amplitude analysis was done while performing lumbar infusion (72,135). Regardless of measurement route, the predictive value of pressure amplitudes in INPH needs to be confirmed in a randomized study.

The underestimation of ICP amplitudes was less at higher pressures. This may in part be explained by difference in measurement systems, where lumbar measurements with a fluid catheter system amplify overtones of the frequency spectra, and more so at larger amplitudes. However, relying on amplitudes derived from the fundamental frequency will underestimate true amplitudes. Even though measurement system discrepancy increased at higher amplitudes, the differences were small compared to the magnitude of the measurement, and may not have clinical relevance.
The limitations and strengths of paper I also hold for paper II, i.e. low number of patients, but many measurements over a wide range of ICPs in each patient. One limitation may be that the baseline ICP and ICP amplitudes of paper II were large compared to other studies on INPH (49,72), and difference between lumbar and ICP amplitudes may be larger at lower ICPs. A general limitation with the method of amplitude measurements the lumbar route, may be influence from physiological factors such as spinal stenosis. However, this needs to be researched.

Lumbar puncture is routine in preoperative assessment of INPH patients. Intracranial pressure pulse amplitudes are reflected in lumbar space, which opens up for studies of the predictive test described by Eide (72,74). Also, following the results of paper II, our research group has described another interesting application of lumbar CSF pressure amplitude measurements. If large preoperative intracranial pressure amplitudes are predictive of outcome after shunt surgery in INPH, and the curative effect of the shunt is by lowering amplitudes, then there should be an association between amplitude reduction and outcome. Accordingly it was demonstrated that amplitude reduction measured through the lumbar route was significantly larger among patients who improved in gait (responders) after shunt surgery, than non-responders (49). However, this needs to be confirmed in a larger study. In the cognitive domain Foss et al. found that INPH shunt responders had larger preoperative ICP amplitudes (98). It remains to be examined if amplitude reductions are associated to improvements also in the cognitive domain. The effect may be even more pronounced by combined measures.

The potential for amplitude reduction from shunt surgery differ between patients (21). The idea of preoperatively assessing the potential for amplitude reduction opens up for interesting predictive tests. Especially important in a field, where the largest challenge is excluding patients who will not benefit from shunt surgery (65).

The prospect of lumbar pulse pressure measurements in INPH seems promising. But to study their clinical impact, accessible and sensitive outcome measures are needed.

**CoGNIT development**

When setting out developing the computer test the concerns were manifold. Would elderly, potentially non computer-skilled persons be testable? How to
select appropriate tests? Also, the purely technical aspects were of concern, 
e.g. what computer language is suitable?

The CoGNIT battery was evaluated with regards to completion rates of tests, 
validity, reliability, discriminative ability and ability to detect cognitive 
 improvement after shunt surgery, discussed in the following paragraphs.

Test evaluation

The technical solution of using Java and Flash programming languages was 
good. Java is a widely used programing language and programs run on any 
computer. In the Flash language development of user interfaces and 
amimations are very easy. When starting the neuropsychological test battery, 
the Java program, starts the Flash application; receive data from the tests 
(and the microphone), and store results to a file. This technical solution 
makes for easy development, fast translation of tests into other languages 
but also portability to different computer platforms (e.g. Apple OS X or, 
Microsoft Windows). It is also clear that the touch screen interface was a 
good solution, providing a natural way of interacting with the computer. It 
has also been proposed as the best input modality for testing elderly (108).

Completion rates

Completion rates were good. The healthy elderly completed all tests. 
Hydrocephalus patients showed a preoperative completion rate around 80 
%, and for individual tests the completion rates were in parity with 
conventional neuropsychological testing. The tests that were not always 
completed were the Trail making test B, Stroop incongruent colors and the 
Four-finger tapping test. Trail making test B was regarded as failed if the 
patient gave up, whereas in the Stroop test, the test was failed if more than 
50 % errors. In paper IV, this criterion was loosened, and only 10 correct 
answers were enough to gain a score on the test. This was in concordance 
with other studies using the Stroop test in INPH (82). Still, completion of the 
Stroop test was around 70 %, which is in line with other studies on INPH 
(82). Trail making test B was completed by 75 % which is far more than the 
38 % completion rate seen in the largest study of neuropsychology in INPH 
to date (79). It should be noted that also a failed test provides information 
about patient performance on a test. This may be more so in the executive 
function domain where there is a component of "getting it". In the Four-
finger tapping test many patients started out correctly but soon reverted to tapping with only the index finger. This might be relieved by adding an extra instruction, reminding that four fingers should be used. However, no data confirming this are present at the moment. Still, completion rates for CoGNIT can be regarded as satisfactory. Importantly, completion rates were unaffected by reported computer knowledge.

**Validity**

The implemented tests were valid, in the sense that they correlated significantly with their conventional paper-and-pen versions. The magnitudes of correlations are dependent on range, but also on reliability (measurement error) in both the computerized and conventional tests. Hence, because of the poor reliability, the Figure-copy test showed a weak correlation in *paper III* and was removed in *paper IV*.

Divergent validity was established between most of the tests. Not surprisingly there were correlations between tests in the same cognitive domain e.g. Trail making test A and Stroop congruent (psychomotor speed), Trail making test B and Stroop incongruent (executive functions) and Ten-word list sum of three trials and Delayed recall and recognition (memory). Also, there were correlations between tests with a strong motor component. Neuropsychological tests are seldom a pure measure of a cognitive domain or ability. In all test with touch screen interface there will be some "motor contamination". A similar problem exist in many conventional neuropsychological tests e.g. in the Trail making tests. Though correlation between tests, it was noted that median improvement after surgery in the Stroop incongruent test was double that of the Stroop congruent test (tests with the same motor component), indicating that functions beyond motor performance improve. It has been argued that divergent validity is more important to establish when evaluating computerized tests (105), as e.g. computer knowledge can influence all scores. Reported computer knowledge was only associated with the baseline score of INPH patients in Trail making test A in *paper IV*. A finding that was not replicated in either healthy or INPH patients of *paper III*. 
Reliability

In paper III it was shown that 8 of 10 the implemented tests had good test-retest correlations ($r > 0.7$). The Ten word list ($r=0.67$) and the Figure copy test ($r=0.57$) were the tests with moderate performance with this regard. Reliability is a measure of error in a score and it can be shown that correlations between test and retest estimate this error (123). However, as any correlation, this number is largely dependent on the range of measurements and test-retest correlations cannot be compared between groups. An absolute reliability measure such as the standard error of measurement (SEm) is less affected by the population under study (136). In this study the SEm was calculated as:

$$SEm = S_x \sqrt{1-r}$$

where $S_x$ is the standard deviation of the observed scores and $r$ the test-retest correlation. The SEm allows for making statements about the confidence level of individual scores. An obtained score is approximately within one SEm with 68% confidence. It is clear that the SEm is less influenced by the variation of scores and thus populations. An increased variability of scores will increase $S_x$ but will be compensated by an increased $r$.

After paper III, it was clear that the Figure-copy test showed poor reliability. Probably performance was hampered because of a dichotomous score, where a difference from pass to fail has a large impact. Also, in this test the result has to be judged by a human, which makes for the traditional errors when human judgments are involved namely intra and interrater variability. This test was therefore left out of CoGNIT. Of the remaining tests it was shown that the test with least test-retest reliability (Ten-word list, $r=0.67$), with great confidence separated INPH from healthy, even when measurement error was accounted for. This is an example of relating measurement errors to "analytical goals", i.e. the implication of the measurement errors on inferences from a test (136). The subtests of CoGNIT were reliable enough to be evaluated for clinical usefulness.
**Baseline impairment and cognitive change**

Pathology in INPH is related to dysfunction in subcortical white matter with the largest deficit closest to the ventricles (137). Thus, impairments in the rich connectivity from basal ganglia to frontal brain areas, connected via so-called basal ganglia-thalamocortical loops, has been hypothesized to cause the observed pattern of cognitive deficits (99,121,138-140). In the context of INPH their disruption would be possible at several sites, at the striatum, most likely at the caudate nucleus close to the lateral ventricles, the thalamus, adjacent to the third ventricle and the periventricular white matter. A finding in INPH, potentially affecting several cognitive domains is that many INPH patients show impaired and reversible level of wakefulness, that may be due to disturbance of the brain arousal system (141). A radiological finding in INPH is reduced hippocampal volume and regional cerebral blood flow, though not to the extent seen in Alzheimer’s (141,142). This would explain the relatively mild memory deficiency in INPH (87).

CoGNIT detected significant impairments in all tests in untreated INPH patients in both papers III and IV, see figure 13. Looking at individual data the interquartile ranges of healthy and INPH scores were non-overlapping in all tests but a slight overlap in Trail making test B of paper III. Overall cognitive impairment reflected by the summed Z-scores in paper IV showed that only one INPH patient scored just above the 5th percentile of healthy. The ability for CoGNIT to detect cognitive impairment of INPH patients must be regarded as good. Patients performed worst in the Delayed recall and Stroop incongruent tests and best in the Delayed recognition test, corroborating previous studies on conventional neuropsychological testing in INPH (62,81). CoGNIT can thus be regarded to capture the "finger print" of cognitive impairments present in INPH patients. Also, the effect of vascular risk factors on neuropsychological performance in INPH has been described before (81,99).

In paper IV, it was shown that CoGNIT was sensitive to detect post shunt improvements in attention/psychomotor speed, manual dexterity, memory and executive functions. No ceiling effects were observed in any test. Some tests did not show any improvement on the group level (Trail making tests and, the Delayed recognition test). Notably the Trail making test B was not found to improve on a group level in the largest INPH cohort studied to date (79). Post-shunt improvement in the domains of CoGNIT has been described before (62,79,82). The magnitude and number of patients who improve after shunt surgery differ between studies for a number of reasons. Firstly,
different neuropsychological tests and criteria for cognitive change are used in different studies. Secondly, the time between neuropsychological evaluation and shunt surgery likely vary between cohorts. Andrén et al. showed that neuropsychological performance merely reached to baseline values when surgery was delayed by more than 6 months (61). This delay is rarely reported in studies and may be a large factor affecting results. In paper IV, the mean pre-operative test delay was 5.6 months, probably hampering results to a bit. Thirdly, centres have different criteria for shunting, and may have different success rates with regards to this. Regardless, the sensitivity of CoGNIT to detect cognitive impairment in INPH patients seems good.

There are several intricacies with regards to measurement of cognitive change. Ideally each patient should be compared with an untreated clone. As this is not possible other methods must be used. In this thesis patients were compared to the measurement error of the test. This is at least theoretically stable across groups. Others have related baseline impairment and improvements after shunt surgery to the standard deviation of a control group, which makes for a more floating interpretation, as biases can be introduced by i.e. different ways of recruiting controls.

Limitations

Limitations in the evaluation of CoGNIT include a relatively small number of healthy controls that may hamper generalizability. Also, the small number of INPH patients in paper IV, is limiting the power to detect post shunt improvement.

The general criticism of computerized neuropsychological testing is that of failure to show validity, and often lacking normative and error data. Also the field is suffering from poorly designed computer-patient interfaces and failure to show influence by e.g. computer knowledge (105,107). In the development and evaluation of CoGNIT these issues were addressed.

It must be stated that testing with CoGNIT only provides data collection (and a written report immediately after ending the testing session). Skilled personnel must perform interpretation and diagnostics. Also, computerized neuropsychological testing does not provide the qualitative data from an evaluation by a neuropsychologist. Of course, before drawing conclusions about higher cognitive functions, lower functions must be assured (e.g. arousal). However, this is not unique for computerized testing.
**Standardization in INPH**

A great leap forward in the field of INPH was the international guidelines for diagnosing the disease published in 2005 (143). A standard definition of the disease makes for comparability of patient material between studies. In the measurement of outcome however, there are a plethora of methods. Several INPH specific scales have been proposed, but none widely adopted (25,144-146). A promising development in this field, is the new INPH scale put forward by Hellström et al. (100). The neuropsychological part includes an average of normed scores of the Stroop tests, RAVLT and the Grooved pegboard tests. CoGNIT includes tests tapping in to the same cognitive domains, but is more extensive, and does not require special training to administer. The MMSE is often included in many studies on INPH and makes for comparison. It is practical as various health workers can administer it, but lacks sensitivity in INPH (76).

**Future directions**

Having a standardized cognitive test that can be administered by e.g. nurses will be practical and if adopted aid in comparability and cooperation between centres managing INPH patients. A perquisite is that CoGNIT is translated to other languages. Today, there are translations into Danish, American English and Swedish. As of today German seems to be the next language of translation. There is a bit of effort attached to this, as norms may differ between countries. Also, translation may to some degree alter validity and reliability of the tests.

The technical solution of CoGNIT test development makes for easy adding and studying new neuropsychological tests, potentially advancing the field of neuropsychological testing in INPH.

A next step could be porting CoGNIT to iPad. An iPad version would be easily distributed and potentially increasing adoption.

The accessibility of CoGNIT may aid the longitudinal management of INPH patients, where a sudden impairment may indicate other evaluations for e.g. shunt dysfunction.
There are few studies that have looked at the utility of cognitive testing before and after CSF drainage tests (69,86,147), and results are mixed. It would probably not be practical to administer the full CoGNIT battery pre and post e.g. a spinal tap. But a subset with the most responsive tests may be useful. These tests may provide data in patients who are not able to walk, where a standard test protocol with gait tests would give no information. With retesting over such a short interval as 24 or 72 hours, practice effects need to be explored and accounted for. A new research protocol is already under way with regards to this.

An interesting research path lead from the finding that gait improvers show larger CSF-pressure amplitude reductions after shunt surgery. Can the findings be replicated in the cognitive domain or combined measures?

Comorbidities are common in INPH (53). CoGNIT is a focused battery for INPH, and the usefulness for differential diagnostics is probably limited. However, there might be clues, as e.g. a very large memory impairment including a impaired Delayed recognition test might give a suspicion of a concomitant Alzheimer’s disease (AD). Also the potential for cognitive recovery may be hampered by comorbidities such as AD (86). However, this is yet to be researched.
CONCLUSIONS

Ia. The pulsatility index (PI) derived from the middle cerebral artery was associated with intracranial pressure (ICP).

Ib. The prediction interval for ICP from PI data was too wide to be of clinical use. The large difference was explained by individual differences in the PI-ICP relation.

IIa. Intracranial pressure waves were measurable through lumbar puncture. There was a difference in amplitudes from ICP and lumbar measurements at resting pressures, likely due to dampening of pulse waves in the spinal canal.

IIb. Lumbar pressure amplitude measurement with fluid catheter is an alternative to direct ICP measurement with a catheter tip sensor, but the thresholds for what should be interpreted as elevated amplitudes need to be adjusted.

III. The computerized neuropsychological test battery (CoGNIT) shows satisfactory test-retest reliability in healthy elderly and validity for cognitive impaired patients. Additionally, health workers without special training can administer it.

IV. CoGNIT show good completion rates in INPH patients. CoGNIT is sensitive to detect the preoperative cognitive impairment and changes after shunt surgery in INPH patients.
ACKNOWLEDGEMENTS

I want to express my sincere gratitude to all who have contributed to this thesis. Without you, and modern technology this thesis would never have come true.

Jan Malm, my main supervisor. You have been supportive, patient, available and helpful.

Anders Eklund, co-supervisor. For your dedicated mentorship. You taught me to really look at the data.

Niklas Lenfeldt, co-supervisor. For your help and sincerity. You got me started in research, for which I am truly grateful.

Eva Elgh, co-supervisor in the field of neuropsychology. For your help, friendly and encouraging way.

The always-friendly Lars-Owe Koskinen for good cooperation.

Colleagues in the group, PhD candidates and co-authors, for interesting discussions and lots of fun, Sara Qvarlander, Anders Wåhlin, Khalid Ambarki, Kenneth Anderson and Nina Sundström. Special thanks to Hanna Israelsson for valuable comments on the thesis.

I would like to thank Michael Williams and Cynthia Smith for great hospitality in my visit to Baltimore and for good cooperation with CoGNIT.

Friendly partners cooperating with the CoGNIT in Sweden and Denmark, Göran Leijon, Katarina Laurell, Bo Traberg-Kristensen and Helle Vibe Krog for providing the Danish translation of the test.

Sonja Edvinsson, Jörgen Andersson, Hanna Ackelind, Carina Gillewård, Marie Bäckström and Kerstin Gruffman at the neurological department in Umeå. Special thanks to always-cheerful Kristin Nyman, a great working companion. Without you, this would be a thesis without any data.

My employer at Department of internal medicine at Karlskrona Hospital for patience and funding. Colleagues at the Department for internal medicine. Especially my skilled colleagues in the neurology department: Marcus Svennerud, Håkan Johansson, Malgorzata Losiak, Björn Servin, Evelina
Kask-Ogenblad, Lars-Erik Sjöberg, Lars Brattström, Maria P Idbrant and Carina Gunnarson. Also, the skilled staff at 58:an.

Vetenskapliga rådet and Kompetenscentrum at Karlsrona Hospital, for funding, a place to work and friendly colleagues. Alzheimerfonden, Kempefonden, Forskningsfonden för Klinisk Neurovetenskap vid Norrlands Universitetssjukhus for funding.

My friends. Especially Daniel Jinnelöv for valuable help on Java programming and Magnus Ekström for valuable comments on the thesis.

My parents for love and encouragement. You are great. My brother and sister. You are great too.

Fanny, my future wife. For patience, love and our wonderful children Lovisa and Astrid.
REFERENCES


23. Williams B. Simultaneous cerebral and spinal fluid pressure recordings. I. Technique, physiology, and normal results. *Acta


47. Eide PK, Sorteberg W. Preoperative spinal hydrodynamics versus clinical change 1 year after shunt treatment in idiopathic normal


78. Kukull WA, Larson EB, Teri L, Bowen J, McCormick W, Pfanschmidt ML. The Mini-Mental State Examination score and the
79. Solana E, Sahuquillo J, Junqué C, Quintana M, Poca MA. Cognitive
disturbances and neuropsychological changes after surgical
treatment in a cohort of 185 patients with idiopathic normal
May;27(3):304–17.

Validity of the MoCA and MMSE in the detection of MCI and

81. Hellström P, Edsbagge M, Archer T, Tisell M, Tullberg M, Wikkelso
C. The neuropsychology of patients with clinically diagnosed
idiopathic normal pressure hydrocephalus. Neurosurgery. 2007

82. Hellström P, Klinge P, Tans J, Wikkelso C. The neuropsychology of
iNPH: findings and evaluation of tests in the European multicentre

83. Hart RP, Kwentus JA. Psychomotor slowing and subcortical-type
dysfunction in depression. J Neurol Neurosurg Psychatr. 1987

84. Cummings JL. Subcortical dementia. Neuropsychology,
1;149(6):682–97.

85. Filley CM. The behavioral neurology of cerebral white matter.

Hillis AE, et al. Baseline neuropsychological profile and cognitive
response to cerebrospinal fluid shunting for idiopathic normal

87. Ogino A, Kazui H, Miyoshi N, Hashimoto M, Ohkawa S, Tokunaga
pressure hydrocephalus. Dement Geriatr Cogn Disord.

88. Iddon JL, Morgan DJR, Loveday C, Sahakian BJ, Pickard JD.
Neuropsychological profile of young adults with spina bifida with or
without hydrocephalus. J Neurol Neurosurg Psychiatr. 2004 Aug


111. Agren-Wilsson A, Roslin M, Eklund A, Koskinen L-OD, Bergenheim AT, Malm J. Intracerebral microdialysis and CSF hydrodynamics in...


