

Biomarkers for Huntington's Disease: the Impact of New Technologies

report by Lisa J. Bain

Abstract:

With potential new treatments appearing on the horizon, the Hereditary Disease Foundation convened a workshop in February, 2002, of a diverse group of investigators to explore new approaches for finding useful biomarkers. The group discussed the possibility that onset, currently defined by motor symptoms, may be an outmoded concept, and suggested replacing it with rate of regression. New paradigms were discussed for identifying, in both humans and animal models, cognitive and motor impairments that may precede motor symptoms; and that may identify specific neural circuits affected by the disease. Eye movement abnormalities may also be useful in identifying progression of disease, although work in this area has been dormant in recent years. Imaging methods may be useful for quantifying early stages of neurodegeneration and brain atrophy. Participants stressed the need to correlate findings in humans with those in HD mouse models, and to use mice to explore aspects of the disease that are difficult to explore in humans. Novel methods of screening drugs in animals using automated systems generated a great deal of interest.

The need for biomarkers for Huntington's disease becomes more and more pressing as potential new treatments appear on the horizon. Thus, on February 3rd and 4th, 2002, the Hereditary Disease Foundation convened a group of some 20 investigators working in diverse fields including clinical neurology, neuropsychology, behavioral neuroscience, neurogenetics, electrophysiology, basic neuroscience, and engineering to consider new approaches for finding useful biomarkers. With representation from both human and mouse researchers, a major focus of the workshop was to draw parallels between markers identified in humans and in mice; and to find means of transferring paradigms from one species to another.

As is common at HDF workshops, the meeting opened with a visit from an individual affected by HD. Carlos Urrutia and his wife, Karen, poignantly shared their experiences over the past 10 years as Carlos learned through genetic testing that he had inherited the HD gene and then as he began to manifest symptoms a few years later.

Carlos' experience with genetic testing is somewhat unusual. In the United States, only 3-5% of at-risk individuals choose to have their gene status evaluated. Walter Koroshetz noted that before testing became available, in a survey of at-risk individuals, 95% said they would get tested; but when people realized the implications, most have chosen to forego the test.

The group discussed some of the positive and negative aspects of getting tested. Carlos said he is just the type of person who "wants to know things." Karen, however, did not want him to get the test, and when he tested positive she was "completely shocked." Some people choose to have testing because of the implications for their future children. Carlos and Karen adopted children before learning of Carlos' gene status. Insurance is another common concern.

According to Blair Leavitt, in studies of long-term follow up of predictive testing, the incidence of serious adverse reactions is the same regardless of the results of the test. "What I'm struck with over and over is how often a negative test has a catastrophic effect on a patient. Because if you happen to be one of the people who live your life with this sword of Damocles hanging over your head, and you say, 'I don't need to have any stable relationships, I don't need to get an education, I don't need to plan for my future... If you suddenly find out that you have to plan for your life, that's just as devastating as if you lived your life and said, 'I may have this disease but I'm going to be the best person I can be, regardless of whether I'm tested.'"

Biomarkers for Huntington's Disease: the Impact of New Technologies - DRAFT

Page 2

What became clear is the need for physicians to understand that the decision to get tested is a highly personal and individualized one, and that at-risk individuals can be aided by professionals who will explore the risks and benefits. Said Walter Koroshetz, "The only cogent reason to take the test is if the knowledge that you have inherited the HD gene can make your life better." It becomes a much more complicated issue, he added, when it comes to developing treatments, as individuals may wish to know their gene status when treatment options become available.

Why biomarkers are needed

While testing for the presence or absence of the mutant gene may not be of benefit to many people, finding a marker that predicts onset might be more useful. Elizabeth Aylward, co-principal investigator of the PREDICT-HD (Neurobiological Predictors of Huntington's Disease) study explained that one of the goals of that study is to find markers that predict onset. The hope is that predicting onset will allow clinicians to decide who, among presymptomatic individuals, would be suitable candidates for clinical trials. Biomarkers could then be used to monitor whether a drug is working, for instance, by delaying onset. PREDICT-HD is a four-year study funded by the National Institutes of Health and sponsored by the Huntington's Study Group. The study will enroll 425 individuals who are gene positive and 75 individuals who are gene negative; and will follow them for four years. MRIs and neuropsychologic testing will be conducted at baseline and again after 24 months. Using standardized neurologic exams and the UHDRS (Unified Huntington's Disease Rating Scale,) a clinician will determine onset.

Onset is defined clinically by motor symptoms. This is despite the fact that cognitive and psychiatric impairments frequently precede motor impairments; and the observation of atrophy in the basal ganglia as seen on MRI long before clinical onset (perhaps as early as 10 years prior to onset of motor symptoms), said Aylward. Koroshetz noted that using motor symptoms as the criteria for onset results in specificity, but not sensitivity. Moreover, motor and cognitive symptoms may be independent. Steve Dunnett said, "If all your focus is on motor systems because that's the easiest to measure, and you solve it with a drug that completely resolves all the motor symptoms, you've still got the disease. You've still got big problems."

Karim Nader asked whether it would be more useful from a therapeutic perspective to define onset in terms of the earliest deficit, regardless of whether it is cognitive, motor, or emotional. Claude Ghez pointed out that a problem with using behavioral measures as indicators of onset is that behavioral measures, whether motor or cognitive, are subject to enormous sources of variance, including the state of the subject (i.e., fatigue, motivation) the interaction with the investigator, and how the investigator describes the test. Nancy Wexler noted that when testing subjects in Venezuela, investigators are trained to use maximal encouragement as an attempt to standardize the administration of the tests. Maurice Smith said that the motor tests that he would discuss later in the meeting appear to be less affected by motivation of the subject than some other tests of motor function.

Marie-Francoise Chesselet added, "If we keep thinking about measuring dysfunction in the system, we will constantly find measurements that depend on motivation and learning, because that's what the system does; it's involved in motivation and learning. So we need to step back and find another way to look at the system."

The discussion about onset led several participants to suggest abandoning the whole notion of age of onset and replace it by rate of regression. Onset made sense before the gene was discovered, said Aylward. Now genetic testing can tell you whether you have the gene but doesn't tell you whether you have the disease. Assessing the rate of change in symptoms might yield more useful information. In terms of assessing the effectiveness of therapy, said Koroshetz, time is the key constraint. Markers are needed that shorten the time frame to say 'this is the rate of progression we expect and this is the progression we got with therapy.' Using something like total functional capacity (TFC) as a marker is too insensitive because the measurable change is too slow.

Biomarkers, particularly of early events in pathology, are also needed in order to speed the testing of potential treatments in mouse models of HD. Allan Tobin remarked that in the next year there are likely to be scores of compounds that will have been shown to have some efficacy in molecular, cellular, and organismal models of HD. These compounds and analogues of them will need to be tested in mouse models of HD in order to identify the best compounds for human clinical trials.

Behavioral and cognitive measures in humans

Claude Ghez and Marie-Felice Ghilardi described tasks they are using to study motor sequence learning in presymptomatic individuals. In these tasks, subjects reach for a target on a computer screen that appears in a predictable or unpredictable sequence. The task differs from classical SRT (serial reaction time task) in that subjects are told to look for a sequence, identify it and predict it. This forces them to learn the sequence explicitly rather than implicitly and shortens the duration of the task. Normal subjects learn the sequence and anticipate the stimuli, while HD gene carriers wait for the stimuli before reaching for the target. Ghez introduced a variant to determine how much of the reduction in reaction time results from an impairment in the ability to move and how much is from an inability to learn the sequence.

Ghilardi and Ghez found that sequence learning is impaired in presymptomatic HD subjects. Moreover, the impairment is greater when movement is required in addition to sequence learning. According to Ghilardi, "Visual learning performance in presymptomatic HD subjects is significantly correlated with neuropsychological tests of processing efficiency, working memory, as well as shifting and inhibition. Thus, aspects of frontal executive functions may be impaired already in the pre-clinical stages of HD. This may be due to functional abnormalities in striato-frontal processing."¹

With David Eidelberg and Andy Feigin, Ghez and Ghilardi are also conducting imaging studies in subjects who are performing these kinds of tasks. In unaffected people, better performance on these tasks is associated with activation of the dorsolateral prefrontal cortex. In contrast, in HD subjects there appears to be increased activation of the anterior cingulate and decreased activation in the dorsolateral prefrontal cortex, what Ghez called a "compensatory activation of the anterior cingulate."

Behavioral and cognitive measures in mice

¹Ghilardi MF, Feigin A, Mattis P, Silvestri G, Veytsman M, Zgaljardic D, Eidelberg D, Ghez C. Sequence learning in pre-symptomatic Huntington's disease (HD). Society for Neuroscience, 2001.

Biomarkers for Huntington's Disease: the Impact of New Technologies - DRAFT

Page 4

Jenny Morton and Steve Dunnett have been looking at cognition in rats and mice. Dunnett described a swim tank task that they have been using as a measure of conditional learning and which is not dependent on complex motivation effects. Mice are taught a primary rule of swimming toward one end (either the dark or light end) of the tank in order to get out. R6/2 mice can learn that task until quite late in their disease course, with impairments in learning beginning to show up at about 9 weeks of age. However, if you probe that by looking at reversal, e.g., training them to go to the light end and then switching the escape platform to the dark end, the R6/2 mice take longer to learn than do wild type animals. "And what that tells you is that the deficit is not one of motivation or ability to swim or to detect the stimuli, because they can learn the primary rule," said Dunnett.

Morton and Dunnett have also used a second set of tests using a variation of the Morris water maze. The way the test is usually conducted, it is difficult to tease out whether deficits are based on motor, sensory, or learning problems. Controls can be built in for visual and motor problems. Then, by building in a reversal component, you can more selectively assess fronto-striatal deficits. Animals are trained on the first day to go to one position; then moved to a second position on the second day and a third position on the third day. Each day, you see how well they carry over what they have learned the previous day and also, how well the animals adapt to change. Striatal and frontal lesions disrupt the learning. Thus, lesioned animals are usually faster on the first trial because they have less memory of what they learned on the previous day, but then they show less change from trial 1 to trial 2.

Ann Graybiel has taken a different approach to finding behavioral measures of HD. She has been interested in looking at behaviors that are associated with striatal function, reasoning that neurodegeneration in the striatum is a cardinal feature of HD. Stereotypies, or repetitive movements, are common in many neurologic and neuropsychiatric disorders, for example, obsessive compulsive disorder. The basal ganglia is thought to be central to the expression of stereotypies. In Graybiel's lab, she has been using amphetamines and cocaine to induce stereotypy in mice and rats, and then has been looking at the differential activation (of genes) in the neurons in two distinct compartments of the striatum, the striosomes and the matrix. Graybiel has documented a high degree of correlation between the amount of repetitive behavior and the degree to which the striosomes are activated.

Nancy Wexler emphasized how relevant this work is for HD. She told of a woman with HD who has had a difficult time dealing with OCD. For her, this has been the most agonizing part of the disease and according to Wexler, is much more common in people with HD than is generally recognized.

Graybiel is now working with Lucy Brown, looking at the striosomes and matrix in several HD mouse models, including the knock-in, R6/2 (transgenic), and Yamamoto (inducible) mice. Although there are no data yet to report, early indications are that in some of these HD models, the striosomes and matrix are affected differently. Graybiel said she thinks these experiments may help explain the topography of cell loss and shrinkage in HD, and how modifiers may adjust which cells are more affected. Marie-Francoise Chesselet said that she has found that in knock-in mice and post-mortem brain, inclusions appear preferentially in striosome. In other words, the two parts of the striatum do not seem to be equivalent in terms of pathologic load, she said.

Bruce Jenkins added that the dopaminergic circuitry, including cingulate, striatum and accumbens, can be

Biomarkers for Huntington's Disease: the Impact of New Technologies - DRAFT

Page 5

imaged. In response to amphetamine stimulus, monkeys and rats show strong activation in striatum and accumbens, but monkeys have much less cortical activation than rats.

Nader proposed a novel approach that would test for dysfunction in long-term memory. Long-term memory formation requires new protein synthesis, and recent research indicates that the transcription factor, CREB, is involved in activating long term memory. Using a research paradigm called mass vs. space training, one could assess CREB function, said Nader. Mass vs. space training is based on the observation that long-term memory is worse when you have many trials close together (mass) as opposed to spaced out. But if you increase CREB, you can change that. Nader's idea is that if you train people with different intervals, you would see a break point that would be a function of CREB activity. "If there's a compromise in CREB function, it probably won't show up as a memory impairment," he said. "But if you take a look to see in the mass training paradigm where the break point is, it's likely to be shifted to the right. So it's a way of behaviorally assaying for CREB function." Moreover, mass training can be applied to different learning systems, depending on which neural circuits you wish to test. For example, if you use a task that tests striatal circuits, you could see whether CREB function was affected in the striatum.

Eye movement paradigms

Adrian Lasker discussed work he had done ten years ago to look at eye movement difficulties in HD patients. According to Lasker, HD patients and people with other basal ganglia disorders have trouble keeping their eyes fixated on a particular location. They have "intrusive saccades," which means that they take their eyes off the target and then bring it back. Although Lasker did not see this deficit in presymptomatic HD subjects, this work was done before the gene was identified so there is little information about when the deficit begins to show up. However, Elizabeth Aylward said she did imaging studies in the presymptomatic subjects that Lasker tested and did see atrophy, even though these subjects did not have eye movement abnormalities.

According to Chris Kennard, the classic oculomotor abnormality seen in HD patients is slow saccades. However this is probably a late-onset phenomenon related to abnormalities in the midbrain. Of much more interest, he said, would be the assessment of cortico-basal ganglia input into these systems, because these are more likely to be abnormal at an early stage. HD patients also have trouble with anti-saccade tasks (the subject is instructed to look in the opposite direction to that of a suddenly appearing target, so they have both to inhibit a saccade to the target and make a volitional saccade to the mirror image location). HD patients and those with other basal ganglia disorders have trouble inhibiting their saccades, although again it is not known if this occurs in the early stages.

Other tasks look at memory location and memory guided saccades. Clear abnormalities have been seen in Parkinson's disease patients, although it is unclear if this relates to a spatial working memory or motor deficit. Parkinson's disease patients are also poor at predictive tasks. Lasker said he had done predictive tasks with both light and sound in HD patients. Interestingly, patients are better at synchronizing their eye movements with sound than with light, but get worse again when sound and light are given together. Normally a person will synchronize saccades with lights after about 3 or 4 saccades; and will start their saccade before the light changes in order that

they fixate on the light as it comes on. HD patients do not do that.

Kennard's work has mostly been done in people with Parkinson's disease and visuo-spatial neglect syndrome. He uses tasks that look at visual search, spatial working memory, and planning, such as the Tower of London task. This task requires subjects to plan the moves of a group of colored balls from an initial state to a given goal state and has been used to look at planning and performance frontal lobe dysfunction. Kennard's lab is interested to see if by analyzing the actual eye movement visual scan path of the subject, they can understand the cognitive processes at work. Normal subjects use a characteristic strategy to solve the problem, but Parkinson's disease patients are not able to do so. Kennard said he thinks there is the potential for getting a lot more information out of neuropsychological tasks by looking at the eye's movements while performing them.

Nancy Wexler asked if there were eye-movement measures that could be done on her upcoming trip to Venezuela. "We examine hundreds of people," she said. "This would be a goldmine." Currently, they do simple slow pursuit tasks and can pick up abnormalities in people three to five years before onset.

Kennard said it would be difficult to move most of these complicated paradigms to the bedside, although some of the complex saccadic tasks, learning sequence, trains of saccadic locations, and task switching tasks could be useful. Also, the Tower of London task might be worthwhile since it has shown clear abnormalities in Parkinson's disease subjects. Analysis of these tasks takes time, he said, but here are easy, cheap, and non-invasive ways of measuring eye movements that could be done. Pupilography, or looking at the response of the pupil to light or to spatial structures, could also be very interesting and is something that can be studied in both humans and mice.

Motor behavior in humans

Maurice Smith described work being done in Reza Shadmehr's lab on the motor behavior of HD patients, both symptomatic and presymptomatic. Subjects are tested on an arm movement paradigm in which they are supposed to move to a target. A robotic arm gives them a force pulse to push them about 4 or 5 cm. off course. Both symptomatic and presymptomatic HD subjects have trouble correcting for these perturbations. Smith also looked at jerkiness at the end of movements and found that both symptomatic and presymptomatic HD subjects have more jerkiness. Smith's interpretation of these findings is that the ability to correct on line for errors in continuous movements is disturbed early in the course of HD.

Claude Ghez commented that this defect in terminal correction may be the critical problems that affects fishermen in Venezuela who, despite having spent lifetimes doing highly balanced activities on their fishing boats, when affected with HD may fall into the water when standing on a moving platform. This may also be related to postural instability, which Lasker investigated several years ago.

One of the advantages of this paradigm is that, while many behavioral measures are strongly affected by motivation and learning, this type of motor function is less affected. Most of the participants at the workshop also agreed that this paradigm could be transferred to the mouse model and might be useful in testing the effectiveness of drugs.

Correlating behavioral deficits to neuropathology – imaging biomarkers

For behavioral tests to be truly meaningful, they must be fully characterized in terms of the functions tested and the specific brain regions associated with those functions. This means correlating deficits to structural, biochemical and histologic changes. Imaging perhaps yields the most objective and least state-dependent biomarkers. This includes volumetric MRI, nuclear magnetic resonance spectroscopy (NMRS), Positron emission tomography (PET), and functional MRI (fMRI).

The PREDICT study will be gathering volumetric data on a large number of presymptomatic HD subjects to confirm and extend observations Elizabeth Aylward has made in a small number of subjects. In those subjects she found significant atrophy in the basal ganglia about 10 years before estimated onset. Understanding the pattern of cell death and brain atrophy in HD is likely to reveal relevant biomarkers that may be useful in screening drugs in mice as well as, eventually, assessing their effectiveness in humans. Brain atrophy in HD patients and mouse models of HD follows a characteristic medial to ventral pattern. However, it is unclear whether shrinkage is associated with cell death or not. Many mouse models have atrophy without cell death, said Marie-Francoise Chesselet. "It's very likely that what you see is atrophy without cell death in the very early presymptomatic subjects," she said. Blair Leavitt added that there are more accurate ways of looking at atrophy than just looking at a volumetric MRI; you can look at the average volume of neurons within the striatum, for example.

Flint Beal said that in R6/2 mice you definitely see cell shrinkage, and that it probably occurs concurrently with functional defects. Bruce Jenkins has been using NMRS to measure levels of n-acetyl-aspartate (NAA) in different regions of the brain. NAA, found in neurons only, is thought to be a measure of neuronal health. In R6/2 mice there is an exponential drop in NAA levels at about six weeks of age, which is also the time when nuclear inclusions and functional deficits are first seen. In Borchelt mice (transgenic mice with 82 CAG repeats,) the NAA decrease is linear rather than exponential. These mice show selective striatal neuronal dysfunction, whereas cortical neurons are relatively spared. In contrast, in the R6/2 mice, atrophy is seen in both cortical and striatal neurons. In symptomatic HD patients, longitudinal data are not available, but cross-sectional longitudinal data indicates a linear drop in NAA levels.

Beal also noted that his lab showed years ago that interneurons in both the striatum and cortex are selectively spared. Ann Graybiel called this a "major clue," since the embryogenesis of interneurons is completely different from that of other neurons. Beal said that cortical abnormalities are diffuse but biochemically worse in the occipital cortex and in the anterior cingulate.

Blair Leavitt concluded that the subset of cells that are selectively vulnerable in HD are all large projection neurons that rely on trophic support. These cells are thus likely to be inherently sensitive to metabolic insult. The unifying hypothesis, he said, is that polyglutamines affect some element of function to which these most metabolically sensitive neurons are sensitive. Interneurons, in contrast, have local connections and do not rely on distant trophic support for their survival. "Selective vulnerability does give us clues," he said. "This has an inherent biologic meaning to someone who understands the nervous system."

Walter Koroshetz noted that combining advanced MRI techniques offers the opportunity to detect changes over time. Striatal volume change can be detected using morphometric techniques. Neuronal loss is reflected in a decrease in NAA; gliosis is reflected by an increase in choline. "It seems to me if there is one thing directly linked to neurodegeneration, (separate from a lot of the potential secondary effects that occur during this process,) it would be some function of changes in the absolute measure of these neurochemicals in an easily measured structure like the caudate." The most sensitive marker of neurodegeneration in HD, he concluded would be the ratio of NAA to choline multiplied by a candidate volume of the tissue.

Bruce Jenkins said that resolution of NMRS is currently about 5 cubic mm in humans and is better in the mouse. He added that it is much more efficient to acquire data from many voxels at once, rather than a single voxel; however this technique is more sensitive to motion. A colleague in his lab has built an "incredibly good" facial mask to hold patients' heads still, but patients do not like it. However, the technique may be more feasible in presymptomatic subjects who do not have motor disturbances.

Functional MRI has been less widely used to study HD. Jenkins said he has done some work using tasks that required subjects to switch between finger tapping and toe movements. In very mild, early HD, he saw increased cortical activation including recruitment of more precortical areas, which was not seen in normal controls. "I'm a big fan of fMRI for neuroscience but not for following therapy," he said. In other studies of Parkinson's disease patients, fMRI yielded a high failure rate and the techniques are labor intensive both in data collection and analysis. He said that techniques with better reproducibility, such as the cortical thickness and morphometry studies Diana Rosas does, are better neuroimaging tools. While MRI has relatively low signal to noise ratios, it has the ability to examine specific neuronal and biochemical changes that may be missed using conventional MRI. "I believe it forms a great adjunct when combined with other MRI metrics including cortical thickness, white matter volume, and perhaps ventricular size," he said.

Mouse to human; human to mouse

Throughout the workshop, transferring paradigms from mice to humans and vice versa emerged as an important issue. A key advantage of working with mice, said Blair Leavitt, is the ability to do detailed imaging, quantitative pathologic studies, electrophysiology, and cell culture analysis. "We're limited when we try to look at the human to things we can only do in people, but we can actually do the same things in mice and then say, 'so how does this measure we looked at in humans correlate with striatal cell volume or with electrophysiology of striatal cells.'"

Several participants raised concerns about using the R6/2 mouse as a model of HD. Concerns relate to the fact that the R6/2 mouse has such a rapid course of dysfunction and different neuropathologic characteristics. In his follow-up comments to the workshop, Karim Nader asked, "Is there a role in HD research that the R6/2 mouse critically fulfills that other mice do not?" He noted that while the R6/2 mice express only an N-terminal fragment of the huntingtin protein, other mice are available that express the full length htt protein, as is seen in human HD patients

Biomarkers for Huntington's Disease: the Impact of New Technologies - DRAFT

Page 9

Chesselet described the knock-in mice she works with. Unlike the transgenics, which have a portion of the human HD gene added to the normal mouse genome, knock-ins have an expanded polyglutamine repeat inserted into the mouse gene. Chesselet has studied two different mouse lines originally generated in the laboratory of Scott Zeitlin: one has 94 polyglutamine repeats (94Q) and the other has 140 repeats (140Q). Both lines display early hyperactivity, or increased stereotypy (at one month of age in the 140Q mice, two months of age in the 94Q mice), followed by hypoactivity (at 2 months of age in the 140Qs, 4 months of age in the 94Qs). The 94Q mice have microaggregates predominantly in the striosomes and no aggregates in the cortex, while the 140Q mice have more widespread pathology and less regional specificity, which is more similar to humans.

Nevertheless, the R6/2 is the most well characterized of all the available mouse models, and, as Jenny Morton said, "We need to use the R6/2 mice knowing what problems we have and say, 'let's find a test that this mouse can do that gives us things we can measure that we can relate back to humans.'" One advantage of the R6/2 mouse is that it develops measurable deficits by 4 weeks of age and is dead by 16 weeks, making it convenient for doing pilot studies in a relatively quick time frame. Morton's lab sees cognitive deficits even earlier, starting at about 3.5 weeks.

Jamie Heywood, however, questioned the value of measuring cognitive deficits in mice. All that is needed, according to Heywood, is to find a marker that reliably measures onset and end stage. Other than that, he said, "this sounds like a fascinating way to generate some understanding of basic cognitive neuroscience but I'm not sure what the deficits in mice are going to tell us about how to treat the disease."

Morton responded that this approach might be adequate for testing drugs for ALS, but for people with HD, the probability of developing cognitive deficits is high and those deficits need to be assessed before they can be treated. Moreover, she said, she has done experiments in which they were able to improve cognitive function but not motor function in mice. Chris Kennard added that if you ignore cognitive function and only screen for drugs that affect motor function, you might screen 10,000 drugs and miss something that could have had a profound effect on cognitive function. Karim Nader noted that if you find a drug that affects cognitive, motor, and emotional impairments, it would most likely be acting early on in the pathologic pathway. Thus, he said, it adds much more power and precision to look at all three attributes of the disease.

Nancy Wexler commented that "the mice are talking to us. We put in the same protein that we have in our bodies and something happens to the mice so that they can't walk or talk or find the Morris water maze, or it has trouble learning things." What can we learn from the mice that we can take back to humans to understand the disease better? And as we understand more about people, what questions should we be asking in the mice? "I think this is extremely complementary," she concluded.

In fact, observations that have been made in mice have significantly informed clinicians understanding of the human disease. For example, noted Allan Tobin, the observation of weight loss in mice prompted researchers to reexamine previously ignored studies of weight and metabolism changes in people. Jane Paulsen said that in designing the PREDICT study, they were looking for markers of orbital frontal, anterior cingulate, and dorsolateral

circuitry. Based on literature searches and lesion models in animals, they put together tasks that they thought were best for assessing dysfunction of those circuits.

“But we also need to learn from anatomists and pathologists about progression of this disease,” she said. If the tail of the caudate is affected first, better tests are needed to assess dysfunction of that circuitry. This area is difficult to look at with imaging, she said, so they have added some sophisticated tests of visual perception. They have also added a timing task that can be manipulated to determine whether responses are internally or externally generated. People with striatal difficulty appear to be unable to continue to respond predictably when the external cue is withdrawn. The deficit is thus seen in both HD and Parkinson's disease patients. Paulsen said, however, that while tests such as this may not be specific what is important is that they are sensitive. “We aren't worried that we're going to diagnose Parkinson's disease by accident.”

Just as observations made in mice have informed the studies being done in humans, the reverse is also true. For example, Elizabeth Aylward commented that atrophy seen in human MRI scans might be better understood by studying, at a cellular level, the shrinkage of neurons or dendrites in animals. Maurice Smith's defect in terminal correction, which he has observed in humans, might also be probed by studying the ability of mice to walk on a beam that has been perturbed.

Screening drugs

Many of the mouse paradigms discussed during the two-day meeting could be used in screens of potential HD therapeutics. Jamie Heywood described a system his company, the ALS Therapy Development Foundation, is using to screen drugs for effectiveness in treating a mouse model of ALS. Between April, 2001 and February, 2002, 40 drugs have been tested with at least 20 animals per treatment arm. Animals are given the drugs via oral gavage, intraperitoneal injection, subcutaneous pump, intrathecal pump, intracerebroventricular pump, or orally in food and water. Treatment effects are assessed in a specially designed rat cage by 24-hour wheel monitoring. Complex data are extracted: acceleration time, maximal velocity, length of time the animal will stay tracked on an element, total activity, etc. He noted that a brake could be installed on the wheel to see how quickly the mouse responds to braking. Of the 40 drugs screened, three have shown “true significance,” and another three look “interesting.”

Blair Leavitt said his lab is working with a somewhat similar system, which measures total beam crossing per unit time. The system is automated and can be adapted to measure rearing per unit time, dark/light phase, etc. Heywood suggested adding advanced ultrasonic sensors that could pick up grooming.

Heywood said he also thought imaging could be made cost effective. Allan Tobin said he had talked with Alan Koretsky at NIH about performing MRI on many mice concurrently. Koretsky said the problem is not image analysis, but informatics. Bruce Jenkins said he has done some studies with Flint Beal, in which 6 rats were imaged at a time. Dr. Koroshetz added that MR techniques used to measure drug efficacy/action in mice are especially attractive if they can also be used in patients to determine whether the presumed effect is actually achieved at the dose chosen for human trials.

Priorities for future studies

At the end of the workshop, participants were asked to comment on the strategies and experiments they thought were most important to pursue.

- Jamie Heywood's *in vivo* screening model generated a lot of interest. Several participants said they thought the cages he used could easily be adapted for use with HD mice. Jenny Morton noted that cognitive tests could be built into the wheel running tasks; for example, stopping the wheel and then providing a cue such as an auditory signal to see how quickly the mouse hops back on the wheel.
- Heywood also suggested a "neurology exam station," a bedside device that could do automated strength measurements, motor control assessment, eye gaze, etc., in one 5 to 10 minute exam. "It would be a big engineering project but I think it's possible," he said.
- Morton also suggested engaging an engineer to devise systems that would automate other tasks, such as the beam walking task, so that data can be collected that is reliable and easy to collect. Blair Leavitt suggested a ladder task; Heywood mentioned a gait device developed at Jackson that picks up 20 parameters. Leavitt commented that an automated cognitive task would be especially valuable.
- Interest was also high for Maurice Smith's arm movement paradigm. Flint Beal noted that obtaining quantitative data, as those tasks do, offers major advantages over some of the softer clinical scales. Finding mouse tasks that correlate with the arm movement paradigm might also prove valuable as a reliable way to screen drugs.
- Renewed interest was expressed in studying eye movements that might distinguish between symptomatic and presymptomatic HD subjects. Testing subjects in the Venezuela cohort was especially seen as valuable.
- Adrian Lasker advocated finding more complex learning tasks in order to increase the load on presymptomatics and uncover deficits. Ethan Signer urged behavioral people to think harder and come up with better behavioral tests, especially ones that are quantifiable. He suggested teaming up with a graphic designer as an avenue to increase motivation.
- Signer also suggested following up with Karim Nader's suggested experiment to find a behavioral readout for CREB function.
- Marie-Felice Ghilardi said the most important thing is to characterize the underlying anatomy and loops involved in behavioral measurements. She proposed using imaging and electrophysiology in conjunction with behavioral measures.
- Marie-Francoise Chesselet noted that two types of tests are needed in mice: one that is high throughput and highly automated, and another that is transposable from mice to humans.
- Jenny Morton suggested going back and looking at the histopathology associated with atrophy.
- Walter Koroshetz advocated study of total caudate context of NAA/choline using chemical shift MR imaging for its potential to detect neurodegeneration over months.
- Koroshetz also suggested developing fast screening of drugs using the fly HD model as is being done for Parkinson's disease.

Participants

Elizabeth Aylward
University of Washington

Lisa Bain
Science and Medical Writer

Flint Beal
Cornell University Medical College

John Chaplin
SUNY Health Science Center

Marie-Françoise Chesselet
UCLA School of Medicine

Steve Dunnett
Cardiff University

Claude Ghez
Columbia University

Maria-Felice Ghilardi
Columbia University

Ann Graybiel
Massachusetts Institute of Technology

James Heywood
ALS Therapy Development Foundation

Russell Jacobs
California Institute of Technology

Bruce Jenkins
Massachusetts General Hospital

Carl Johnson
Hereditary Disease Foundation, Cure HD Initiative

Chris Kennard
Imperial College School of Medicine

Walter Koroshetz
Massachusetts General Hospital

Adrian Lasker
Johns Hopkins University School of Medicine

Blair Leavitt
University of British Columbia

Jenny Morton
University of Cambridge

Karim Nader
McGill University

Jane Paulsen
University of Iowa

Maurice Smith
Johns Hopkins University School of Medicine

Ethan Signer
High Q Foundation

Allan Tobin
UCLA Brain Research Institute and
Hereditary Disease Foundation

Minka van Beuzekom
Hereditary Disease Foundation, Cure HD Initiative

Nancy Wexler
Columbia University and Hereditary Disease Foundation