

22 *Per Bech*

Measurement & organisation in psychopharmacology

Can we begin with how you came into psychiatry and why you moved into the area of measurement so early in your career?

I only studied medicine to become a psychiatrist actually. One of my colleagues from the high school had the same idea but when he came to the psychiatric department as a student he called me and said no it was not really what he wanted and he is now a dermatologist. But I always had the idea and so when I finished my examinations in 1969, I came directly to the Psychochemistry Institute in Copenhagen under Professor Rafaelseon. Actually at the same time the clinical professor of the department of psychiatry, Villars Lunn, wanted me to measure a more philosophical subject namely the experience of time in depressive patients. So I went into two areas of research at the same time. The first area, one was to measure the experience of time in depressed patients and there I needed some standardisation about what is a depression, endogenous versus non-endogenous, and on severity because the literature on this subject indicated that altered time experience may be a measure of severity rather than diagnosis.

The second area was that Ole Rafaelseon at that time in the Psychochemistry Institute wanted to look at the effect of Cannabis – this was in 1969. This research project was delayed because we decided to use as experimental subjects males in the military but the Minister of Defence did not think he could allow that because of a fear at that time that Cannabis could perhaps give long term damages. Another Minister finally said that we could use conscientious objectors. The Minister of Domestic Affairs under whom The Civil Defence belonged was very eager to obtain knowledge of the effect of cannabis on car-driving. He gave us permission to study 10 young men. So we made a small study on the acute effect of tetrahydrocannabinol on driving. We had at that time what we thought was a measurement of the blood level so it was a study to measure which blood level would be dangerous for car driving.

Now because of my interest in time experience which is also of interest with cannabis, I combined these areas of research. At this time I picked up the Hamilton Scale and Beck's Depression Inventory so I used those two scales and did a validation study. and in the Cannabis study I introduced the experi-

500 The Psychopharmacologists II

ence of time as one of the extra psychological dimension things apart from car driving behaviour in a car simulator – the Cannabis part has been published in Nature and Science. We used Cannabis as a cake but in a US study they smoked it. They found no effect on car driving but the content of THCA in our batch was tested one year after and there was still the same amount but in the US study there was no active principle left. We showed, compared to alcohol, that if you went up to an equivalent level in the blood (to one per thousand) the subject could not comply with the instructions and so on. I also described that in a Cannabis intoxication you can have a feeling as if time has lost dimension but if you ask for an objective estimate it is good enough was also impaired. In the depressed patient it was more a subjective feeling. They could estimate time intervals as good as others objectively but they had the feeling that time passed on slowly. In the Cannabis study I realised that to measure the clinical experience under Cannabis intoxication you need to have the same scale. Some scales measured positive euphoria (????), others negative euphoria and it was difficult to compare scales. So it was because of this I went into psychometrics and scale construction. In Denmark clinical psychologists in the 60's and 70's were against the use of questionnaires or rating scales. They used in their daily practice the Rorschach test and Murray's Thematic Apperception Test. However, both these tests have very low reliability and no validity in depressive disorders. Danish psychiatrists had no experience either with rating scales. I am really self-taught.

In the depression study I asked had two of the best clinicians in the department to make have a 15 minutes interview talk with the patient and give a mark from 0 – no depression – to 10 – the most severe depression. They had never seen the Hamilton Scale and it was an unstructured interview. Half of the patients they interviewed together and the other half one of them did first and the other in another room later. With the Hamilton Scale I got two Hamilton raters to do the same – they saw the patients at different locations but at the same time of the day in the morning. And then the patient filled out the Beck scale. There was only 24 patients at baseline and four weeks later but when I sat there to look at the items and so on I realised it was difficult to correlation coefficients because they were very dependent on the dispersion. When I had collected 24 depressed patients and reassessed them four weeks later, Ole Rafaelsen asked me to stop the investigation to evaluate if such scales had any meaning in clinical research. Just like the psychologists, he was very sceptical. Although the number of patients was rather small I had the advantage compared to other such studies that the inter-rater reliability both of the global depression scale and of the Hamilton scales was high. When I correlated the scales with the global ratings I realized how difficult it is to interpret the coefficients because they are so dependent on the dispersion. Thus, one single outlier patient could change the coefficient dramatically but I found some methods that compensate for that. I also found out that with the Hamilton Scale when you come up to what experienced psychiatrists will call a score of 5 or 6 on the global scale from 0 to 10 you will score 22 on the Hamilton but when

the experienced psychiatrists go up to a 9 on the global the Hamilton only goes up 25. So I looked for which items the experienced psychiatrist had used and I came down to six items. At that time Max Hamilton said to me 'the best scale for your information is not my own scale, it is the Cronholm-Ottoson scale'. This was designed by these two Swedish psychiatrists who looked at change in depression symptoms during ECT and so I consulted looked up their scale and added some of their items on psychomotor retardation to the ones I already had and that is where the Bech depression scale came from.

This was the same year that Montgomery and Asberg came out with their scale and essentially there is not much difference between the two scales in principle. But as Rafaelson, who now was convinced that the future of clinical psychiatry was in rating scales, told me 'Per Bech, you have to go around the world to sell your scale there is a lot of politics in such a scale'. I refused - did not do that at that time. So I still think it is a good scale but it has not been used as much as it could have been - the Hamilton scale is still the one people refer to. So in summary, in the beginning I went into how to measure things in my cannabis work and in depression and then later I took the same approach to mania and anxiety and personality issues. Scaling problems became my passion in clinical research.

When the Hamilton scale came out first, people were not happy with the idea that you would reduce the richness of clinical reality to this kind of scale. Max Hamilton, when asked about this said that there was some truth to this but that very often the rich clinical reality is not actually assessed all that well whereas at least if people use a scale you can be sure that certain things that should be asked will be asked. But scales like this do introduce a standardisation into psychiatry which is possibly like the standardisation that was introduced into the manufacture of motor cars - it has its good and its bad points. What do you think?

My scale and the Montgomery-Asberg scale along with the Hamilton were picked up in the 1970s and then came DSM-III. Now most psychiatrists agree they had the same idea - to have around 10 symptoms defining, for example, depression. So even in diagnostic systems, the development has been to screen for some symptoms - around 10 - and then you say if there is 5 or 6, you have a diagnosis. so I think that this kind of standardisation had even gone into the diagnostic systems. The only difference is that with the scales you can measure improvement but I think that essentially we train young psychiatrists to look at the same universe of items for depression and I think that standardisation in outcome measurements will be the next thing. Just as with cars you need them to be reliable, so for quality of care you need some kind of standardisation. Outcome is, of course, the most crucial variable in therapy and we have developed the major depression scale both in a DSM-IV and ICD-10 version.

But of course On the other hand, it is common now to ask what is happening to the field of psychiatry? People seem to want something else - a more comprehensive view of a person? This is where the quality of life comes into a field - this offers a more holistic view of the person and measures both what

502 The Psychopharmacologists II

is positive and negative, whereas the Hamilton scale of course goes only for the negative symptoms of depression. Actually I always say to young doctors coming to my clinic, in principle psychiatry is a very easy thing we have such a small number of disorders – depression, anxiety, mania, schizophrenia, dementia and one or two others. Essentially in the daily clinic when we have our strategies for treatment, it is really a small number of dimensions we work from. But coming to the individual depressed patient we need to know more about the patient to give them the best treatment – which comes back to the clinical arts of medicine. But in the emergency department there are only a few things we need to make a diagnosis and there has to be some structure in such situations. I also know a lot of family doctors who like to have the Hamilton on their tables just to screen for depression.

Can I ask you about Ole Rafaelsen? Quite a few people that I have interviewed have referred to him. Can you tell me a bit more about who he was, where he came from and what you think was important about him in the scheme of things?

He was originally on his way to a career in internal medicine. His scientific papers were on diabetes. He had the hypothesis that insulin would not get into the cerebrospinal fluid. Now there were some papers from 1922 or 23 from Vienna that you could use lithium for diabetes and his basic interest when he went from diabetes into psychiatry was essentially manic depressive illness – it was his really only interest in psychiatry actually – and there he looked at lithium and thought At that time the University of Aarhus was the best place in Denmark for studying diabetes. He had to supplement his clinical education with psychiatry and he worked with Erik Strömberg and Mogens Schou in Risskov. Here he became very interested in lithium which actually before the development of insulin was used for diabetes in Vienna. He had the idea that from a genetic point of view you could either have diabetes or manic depressive illness. He tried to push people to look at the how few many patients have both diabetes and manic depressive illness but he was not an epidemiologist. He was really an internist in medicine and who wanted to come closer to the brain. I worked with him there at the beginning of the 70s, doing lumbar punctures and things like that to measure serotonin but I said that even for that we have to measure the severity of the condition and improve our diagnoses. When he finally accepted this and saw that the work he had been doing wouldn't lead to a Nobel prize, he became more concerned about scales and it was then that he invited Max Hamilton to our clinic.

But he was never a natural psychiatrist. Rafaelsen's own psychiatric activities in neuropsychiatry were rather limited. Every time he went to an international meeting and heard that you should measure this or that metabolite in cerebrospinal fluid, he came back and said we should try to measure it. but we could never. We had, however, often difficulties to replicate others findings. We were more among those who could not replicate others work. There was no new thinking – it was more a case of trying to test others hypothesis. He was in a sense an old medical internist. He wanted to measure electrolyte things in

the csf. But lithium has so many different actions and nobody knew what was the most relevant – he thought it was only a matter of time but it was not so and he never had a breakthrough paper people can refer to and his institute is no longer there. But he was a great chairman of meetings and he was a stern man when it came to criticising our manuscripts, which pushed us younger people on. Above all, he created a group of young biological psychiatrists around him in the late 70s, among them, apart from myself, Bolwig, Hemmingsen, Kragh-Sørensen, Rosenberg and Vestergaard. They are today all of them important professors of psychiatry in Denmark.

When you were working in the manic depressive area did you have much contact with Mogens Schou?

Unfortunately not, partly because Ole Rafaelsen in a way belonged to a generation between Mogens Schou and myself. Already in 1953 Mogens Schou and Strömngren made the first placebo-controlled trial in clinical psychopharmacology that showed the superiority of lithium in the acute therapy of mania. Schou and Strömngren shifted the weight of Danish psychiatry towards biological psychiatry and persuaded such bright medical doctors as Ole Rafaelsen to go into psychiatry. Rafaelsen, then, as I just mentioned, in the 70's inspired my own generation to continue the Danish tradition in biological psychiatry.

Concerning the use of lithium in Denmark some interesting differences emerged when Rafaelsen moved from Aarhus to Copenhagen. Thus, Rafaelsen felt that lithium carbonate was the best lithium salt whereas Mogens Schou thought it was lithium citrate. In other words, in Aarhus, lithium citrate is used most often and in Copenhagen we use lithium carbonate – which is funny in such a small country.

Then in the mid-70s when there came some cases from Aarhus of people who were intoxicated with lithium with kidney problems and, Mogens Schou was very concerned about whether he had to stop all Lithium treatment. He approached Ole Rafaelsen and they tried to look at it together. They found out that it was not the case and indeed that they could go down to a lower level of lithium daily. At that time we guided the dose to used over 0.80 nmol/l lithium in the blood but now we use 0.50 and don't see any problems.

In 1976 I published a paper which I had actually started on two years before one summer when Rafaelsen was away. There was a secretary there doing nothing, so I said to her let us go through all records in this hospital from when the first patient received Lithium. The inclusion criteria was that patients should have had Lithium for at least two years. The first patient who took lithium in Copenhagen was the 23rd of December 1959. I was able to find close to 80 patients and I made up a questionnaire about whether they thought it had had some effect on their illness and one of my colleagues was interested in weight gain and compliance. I was actually interested in personality – I wanted to replicate the findings of Carlo Perris that unipolars are more neurotic than bipolars. When Rafaelsen came back and the secretary had told him about all the things we had done, he said to me don't start such projects

504 The Psychopharmacologists II

without permission. At that time we had no ethical committees and anyway such an investigation could be carried out within an interval equal to the period of latency to get approvals from ethical committees you must not do that without asking me. Two days later he came to me and pointed out that I we had found that in our the so-called lithium clinic only 59% received lithium which was one of the results of my small study. This made him very happy because it indicated that we used drugs other than lithium – so it was not a lithium clinic it was a mood clinic. firstly. This was at the time when there was some question about Mogens Schou and Ole Rafaelsen discussed, as I just mentioned, whether treatment with lithium should be stopped.

The next point was that of those who had been treated for so many years with Lithium, only 35% received had Lithium alone. Some patients with a greater tendency to mania had a small dose of Hhaloperidol with it while for those with a tendency to depression, at that time we used received Aamitriptyline, which was Rafaelsen's favourite among the tricyclics. We published my small study as a modern management study of lithium use. Eighty percent of the patients felt that they were helped by liLithium. Those who had stopped had done so mostly because of weight gain. Poul Baastrup who collaborated with Mogens Schou when lithium was introduced called me when he read the paper we had published in Acta Scandinavica and said that in his clinic much more than 35% received lithium alone but 5 years later he called me again and said whenthat now that he had checked it even less than 35% received Lithium alone. I think that this is still the case – in a few patients it is used as the only drug and it is of some help but often it is not enough on its own. But that kind of study of course was not what Ole Rafaelsen meant for his Psychochemistry Institute. For my part I moved more and more towards such clinical outcome studies. In 1980, we created what we called the Danish University Antidepressant Group (DUAG).

Yes. Can I ask you about who actually created that and why?

The main person was Lars Gram. He had been working in the Psycho-chemistry Institute but he went over to clinical pharmacology and he is now a professor in Odense in clinical pharmacology. He went to Pittsburgh and worked there for some years mainly on plasma levels of tricyclic antidepressants. met Ellen Frank and David Kupfer, in the late 70s. When he came back he became professor in Odense. At that time I had to go to other hospitals and there together In the late 70's with Lars Gram, Niels Reisby and I we did two studies – one on the plasma level effects of imipramine and the other on chlorimipramine and showed that, if you used six items on the final Hamilton scale as outcome measuresore six weeks later, the more imipramine there was in the blood the more you lowered the Hamilton scale scoreore. If you used the whole Hamilton scale some side-effects also were measured. I actually made a hypothesis which was published in 1981 essentially about the Hamilton scale and the Newcastle scale using this as an indicator for external validity because

the more endogenous the disorder was the more clear-cut/clearcut this plasma level effect relationship was. In another study we These studies also showed that 225 mgs of imipramine was equivalent to 150mgs clomipramine.

Then came the Lundbeck drug citalopram – the first SSRI synthesized in 1971 – and it was discussed whether we should try to compare that to clomipramine as the standard. It was still uncertain whether we could dose up to a certain plasma level, so we gave a standard dose of 150mgs and we asked if we could make our own protocol. So Rafaelson's department, Odense and Aarhus merged together with the Rijkshospital where I became head of the department and where I still am. So we started in 1981. Then in the same manner from 81 to 83 we looked at paroxetine and then moclobemide. When that was finished I wanted to do a study with a combination of mianserin and clomipramine but Lars Gram asked which dose of clomipramine should we do and suggested we make it a dose finding study. Ciba-Geigy was interested in doing this, so in the last study we have done we used from 20mgs up to 225mgs clomipramine to see if there was a dose in fact it has never been done with clomipramine. I was working still with Rafaelsen in 1980 when DUAG was established. Lars Gram at the Department of Psychiatry in Odense was chairman and leading centre. Niels Reisby was then professor of psychiatry in Aarhus after Erik Strömberg. In 1981 I went to Frederiksborg General Hospital, which was in the northern area of greater Copenhagen, while Rafaelsen's Department at Rigshospitalet was in the middle of Copenhagen. So DUAG originally had four departments which, however, has been extended over the years.

The first trial with citalopram versus clomipramine was published in 1986, the second with paroxetine versus clomipramine was published in 1990. Both studies showed that clomipramine was superior to the two SSRI's. Both protocols were drafted by the DUAG committee, who for instance required that we used fixed doses of the drugs because for plasma level-effect trials fixed doses is the most appropriate. Our third DUAG study was moclobemide versus clomipramine, which has been published in 1993, again clomipramine was found most effective.

The protocols of the trials were drafted by DUAG as mentioned, but they were monitored by the drug companies, following Good Clinical Practice. However, our latest, still unpublished DUAG study was monitored by DUAG itself, thereby showing a rather clear independence to the drug companies. The background for the fourth DUAG study was actually my wish to compare clomipramine with mianserin as well as with their combination. Organon was willing to support this study. During the discussion in the DUAG committee the dose of clomipramine in this combination study was problematic and a dose response study with clomipramine was then suggested. Ciba-Geigy was interested in such a study which was lacking in the literature. This, fourth study with dose-finding of clomipramine has now been finished but not published.

506 The Psychopharmacologists II

One of the interesting thing that came out of all of those trials was that Clomipramine seemed to be more potent than anything else. Do you think that is true? All sorts of people have argued why the results of these studies are wrong.

It must have been around 1987 when we had the results from the paroxetine study but instead of saying that we were very good all the other could not see the right thing, I became interested in meta-analysis as a tool to look at disagreement between studies. The outcome of that for instance in the case of citalopram is that there are two trials where citalopram is not as good as the comparator, with clomipramine in one case and maprotiline in the other, in both cases in inpatients. Most of the trials with citalopram have been in out-patients and similarly with fluoxetine there was only one or two trials for in-patients, so I think that there is something about being in-patient.

I have a PhD student, who has tried to compare the DUAG in-patients with trials carried out in Denmark in the same period in out-patients with essentially the same kind of scales and a similar total score on the Hamilton. The DUAG study had a score of 23 and the out-patients 21 – only a 2 points difference. But the differences are on the items for depressed mood and sleep. Patients who end up in hospital don't sleep – it may be they end up in hospital because their family feel they have to control them day and night – so there is that and suicidality and perhaps clomipramine is more powerful to control such things early on.

When I went through the meta-analysis, I found that in some studies they used only the Montgomery-Asberg scale and in others only the Hamilton. From a statistical point of view the most elegant way to go into this problem of using different scales for different studies and to compare different scales was to use a 50% reduction. My first finding was that a 50% reduction equalled what global improvement calls moderate to excellent. Now from other studies carried out in the new drugs it is closer to 45% than 50% so I have recalculated the DUAG study using the criterion of 50% reduction and after 4 weeks less than 40% had received that outcome in the citalopram group while 68% had it in clomipramine. In the paroxetine study, which was a 5-week study there was no difference if you went to the last week – it is as if you have to wait a little bit, it is not early onset of action for the SSRIs, it is more a delayed onset of action. And in the paroxetine study in the protocol we excluded those who did not score below 15 on the Hamilton after 4 weeks, so many paroxetine patients were excluded who could have responded if we had given them a week more I think.

But in a modern hospital you cannot have patients there too long and that is why we started on the mianserin and imipramine studies. You know when you give citalopram, the effect on serotonin comes within the first day but the clinical effect become apparent after six weeks and so people have been interested in downregulation of beta receptors for instance which is one of the parameters which seems to take some time. In an animal study by Dr Kaisler in Copenhagen showed that if you use imipramine and add a small dose of

mianserin, then you could increase the rate of downregulation so we did a study on this. Those patients who were too ill to go into the drug study received Imipramine, in a dose adjusted for the plasma level, and then either placebo randomised or 30mgs of mianserin. There we showed that the combination was better; those who received a combination went from 25 down to 7 in six weeks and those with imipramine alone went only down to 12, which was significant although it was only a very small number. It is whether it is the Alpha-2 blocking effect of mianserin. We have just finished the same study with 20mgs of fluoxetine adding either placebo or 30 mgs of mianserin and we have shown that the combination gives exactly the same response after 4 weeks. In our first study we measured Hamilton every twice a week because we thought there might be a chance of an early onset of action but it seems as if it is more an effect coming after 4 weeks which is relatively powerful. the combination is useful in the daily routine because mianserin combats nausea and improves sleep, so it is a very commonly used combination in Denmark.

Another thing we have done came from the first meta-analysis when I naively sent a letter to the companies and asked them to do could I use the proper Hamilton using only the first 17 items and also take the sleep factor out and this meant I could go through a huge number of trials. During that I got to assess those patients who had continued treatment after the first six weeks and I was impressed that when they had responded and treatment was continued for six or 12 months they don't relapse actually. So we did a study in my clinic for those patients who accepted ECT – they were then randomised to paroxetine or imipramine and placebo. They all as a mean received eleven ECTs and then they were followed six weeks after the last ECT when in those who received ECT and imipramine the Hamilton score was a little bit statistically lower than in those who had received paroxetine but in the next six months only 12% relapsed on paroxetine while 30% relapsed on imipramine and in those on placebo it was 65%. The placebo were however 10 years older and had a lot of co-morbidity and we measured the plasma levels and perhaps the imipramine should have been higher. But anyway I think the SSRIs may play a role in relapse prevention in some way.

With my interest in psychometrics and rating scales, my statistician was always saying to me that I brought patients who are ill who become better but couldn't we look at some patients who were not ill and then became ill. This study gave us the opportunity to say when patients relapse, how do they relapse symptomatically. We found that there was the same structure essentially that anxiety and depressive symptoms came first and then came retardation elements – guilt and suicide came with those who developed a more severe picture. Cognitive changes came relatively early and introversion or lack of contact was one of the first things when they started to relapse. We now therefore have the hypothesis that if you give an SSRI continually you can control those items and if you can do that you will not go on to the more severe symptoms but if you come to the hospital in a full depressive episode you need other receptors and you may need ECT, which was very powerful in our study – I

508 The Psychopharmacologists II

think 90% had a remission even though the Hamilton score was over 30. It must have been around 1987-1988 when the results of our second DUAG study emerged confirming the citalopram study that I asked Lundbeck, Duphar (fluvoxamine) and Lilly (fluoxetine) to do a meta-analysis of all controlled trials with these drugs, against placebo and tricyclics. This idea was accepted by the drug companies. In my letter to these companies I have asked only to use the first 17 items of the Hamilton Scale, including my own factor of melancholia (core symptoms of depression), the sleep and anxiety factor. Among the outcome criteria I asked for a 50% reduction of the 17 item Hamilton Scale from baseline to endpoint which I equalled to a moderate to excellent improvement on the Clinical Global Improvement Scale. To my surprise no difference between citalopram or fluvoxamine and tricyclics was found in such a meta-analysis. With Dr Cialdella from the Department of Pharmacology in Lyon, France, I re-evaluated the citalopram data using an intention to treat approach. Again no difference between citalopram and tricyclics emerged.

However, if in the citalopram data you select only pure inpatients you will find two trials, the DUAG study with clomipramine and a Belgian study with maprotiline. Both trials favoured the older drugs. Hospitalisation seems to be the most important predictor for showing superiority of tricyclics over SSRI's. Hospitalisation in the 1980's meant therapy resistant patients. It is my conviction that clomipramine is superior to amitriptyline, which again is superior to imipramine in therapy resistant depressions.

Of course, ECT is the most powerful treatment of therapy resistant depression. In a study which has just been published in *Acta Psychiatrica Scandinavica* we have in a randomised way combined ECT with imipramine, paroxetine and placebo. During acute therapy the imipramine – ECT combination was slightly but statistically significant better than the other combinations. However, in the six month follow period after the last ECT, paroxetine was better to prevent relapse. Thus, only 12% relapse on paroxetine, 30% on imipramine and 65% on placebo.

From my experience with patients who have participated in the DUAG trials most of them have preferred the SSRI's in the long-term prophylactic treatment.

But in the full-blown depression treated in hospitals combination therapy is important. We have just finished a study with fluoxetine 20 mg daily plus 30 mg mianserin daily compared to fluoxetine 20 mg daily plus a placebo. The results have confirmed that superiority of the combination therapy.

Finally, I can add that one of my PhD-students – Kurt Stage – has compared the DUAG inpatients with trials carried out in Denmark in the same period on out-patients using the same rating scales as DUAG. The item analysis showed that in-patients scored higher on depressed mood and sleep. Patients who end up in hospitals don't sleep and their family feel that it is difficult to control the patient day and night. Benzodiazepines don't work in this case – most of the DUAG patients have received oxazepam. However, clomipramine, mianserin in combination, or ECT works in these patients.

Looking at the work that was done by DUAG, a lot of people all over the world have been impressed at the independence of the work and they say that the psychiatric profession should be doing more studies of this kind – studies that are not driven by the marketing requirements of the pharmaceutical industry. How come you have been able to do it, when others haven't?

In my own department we have now had over 10 years of studies where the outcome is not as clear because the patients with this kind of depression are not responding so well to the new drugs which are perhaps not so powerful but and I think the future of the work is to make prophylactic studies so that each department can treat the acute part in their own way and then after two months they can go put their patients into a prophylactic study. But during these years we have met every few months to make inter-rater reliability and published the results to show that those who have been treating the patients could manage to use both the Newcastle and the Hamilton scale. So we met relatively often during the study and because of this young doctors in Aarhus for instance got to know people in Copenhagen with whom they could later come and train so there was some advantage in it for the young doctors, so we had this ability to recruit young doctors to do the work. I think it is more and more difficult to do studies of that kind in hospitals today in Denmark. If you want to do this kind of study today you go to the private specialists those who work outside hospital psychiatry. The secret of DUAG in my view has been that the committee covered persons who had worked together beforehand – Lars Gram, Ole Rafaelsen, Niels Reisby, Per Kragh-Sørensen, Per Vestergaard and myself. We covered also different aspects: pharmacokinetics, pharmaco-dynamics, psychometrics, clinical experience. Our relation to the drug companies was scientifically good. The first two drugs are developed by Danish companies – Lundbeck and Ferrosan. Moreover, Roche with moclobemide and Ciba-Geigy in the case of clomipramine had excellent medical research units in Copenhagen. During the trials, from 1980 to 1994, we had monthly rating sessions to maintain a high inter-rater reliability of the Hamilton and Newcastle scales for the participating research psychiatrists. Many people think that this training had a high impact on our research finding, i.e. the ability to differentiate between treatments. In this regard, Professor Zitman from the Netherlands published in 1990 a paper showing that no two research centres could be found in his survey where exactly the same version of the Hamilton Scale was used. Max Hamilton, himself, did not accept the American version used by the drug companies but he did accept the DUAG version in 1986.

DUAG is not doing a fifth trial on the acute therapy of inpatients. Over the many years with trials in which many patients did not respond adequately the departments have had difficulties to recruit patients, if not in a combination therapy. In the future DUAG will perform long-term trials.

You have also constructed rating scales for mania, social dysfunction and more recently for aggressive behaviour. Do you want to comment on any of those in particular?

The mania one came first. When I started with Rafaelsen in the institute we could not find a good scale for mania. I asked Hamilton and he said that very

510 The Psychopharmacologists II

few patients in Leeds became manic so I created this one. We had a meeting that Hamilton participated in and he corrected some of the items and then we published it. There is one item on aggression in this scale. Later on colleagues said you should be more concerned about aggression in the manic state. Also the drug company Duphar who had produced eltoprazine approached me and said that they could not find a scale which covers aggression. So I set up a European group and tried to make a scale. We gathered the so-called experts in Europe. As part of my work on scales you find out about things like protocol adjustment analysis – for instance if you invite experts they are often relatively clear of the number of items needed for a scale but when it comes to the weight rate you would give the individual items there you can discuss and discuss. We had some meetings to come up with a list of relevant items then we got into discussions about introverted aggression and outward aggression and also that aggression could be an episode of one minute or shorter whereas in the manic state we have sustained aggression.

For instance, in my first study in mania I used a US scale for the nursing staff where you had 26 items which would be measured in severity and frequency and it was also some depressed items in this manic state scale. Our study showed that some depressed items correlated well with the manic. I had used the same two experienced psychiatrists who had been in my depression studies. They correlated relatively high with each other which was good because one of the problems if you use two experienced psychiatrists they are often not reliable together. Anyway when I showed my results, one of them said to me 'please let us go down to our library' and he turned to Kraepelin's textbooks where Kraepelin noticed that now and then you see a manic patient for a short period, 5 minutes or so, may become depressed and then they go back to the manic state again. as the Bech-Rafaelsen Mania Scale (BRMAS). It is still the most used rating scale in the pharmacological research of mania. In a recent issue of Psychopharmacology, Post and his group at NIMH in Bethesda have published a meta-analysis of anticonvulsant therapy in mania which covers the period 1978 to 1991. Of the specific mania scales BRMAS is the most used but the BPRS (the Brief Psychiatric Rating Scale) was most often used for general psychopathology.

It is interesting that BPRS has more items on aggression than the specific mania scale. Aggression is an important dimension in a manic state. When the drug company Duphar from the Netherlands contacted me for developing a specific aggression scale I was prepared because the BRMAS was limited in this respect. Duphar had produced eltoprazine which was considered to have a specific action on aggression. I asked some of the experts in Europe both in aggression and in rating scales to join me. After some meetings we set up the European Rating Aggression Group (ERAG) and developed the Social Dysfunction and Aggression Scale (SDAS). It was published in 1992.

I have been very happy with this publication because we were able to illustrate the different statistical models for measuring the internal validity of a rating scale. When I first introduced the Rasch analysis into psychiatry in the late

70s nobody really understood the model. So we had the same issue – should we measure aggression for a short period or should we aim at a more long lasting thing. Actually the scale was made so that we could do both things. There was also exactly the same thing as in mania – where the first scale published was for the nursing staff because they of course are so close to the patient. For aggression there were some self-rating scales interestingly enough and then there was this nursing scale. Our approach was that the information in principle should be collected by the nursing staff and the psychiatrists would come and do a work-up together with the staff. One of the conclusions of that study was that today we should not differentiate between a nurse scale and a psychiatric one – it is the same scale and you can do it together or independently.

We published in 1992. I like this publication because when I did my thesis I used what is called latent trait analysis which nobody really understands what it means. There is what is called coefficient of homogeneity and I said to my statistician 'please make a way to illustrate the latent structure analysis so that it correlates to the coefficient of homogeneity because doctors, myself included, like a coefficient. If we know what is the right value to have, we are happy – other things are a little bit too complicated'.

Duphar anyway used the scale before it was fully established – they could not wait apparently. As I understand it, eltoprazine had some effects on severe aggressive episodes but not enough so I think they stopped it. But we have a scale as a result and it is now used in manic studies or in schizophrenia. For example, Clozapine in our country is used more more aggression than...However, most psychiatrists know what correlation coefficients are and most scale developers including Max Hamilton have clearly been influenced by Spearman's common factor theory when measuring the construct validity of their depression scales. It has always been my approach not to rely on coefficients as used in factor analysis, because they are too sensitive to sample selection. My approach has been to have an a priori theory, namely the coherence of items in terms of their hierarchical tapping of information along a dimension of severity of depression. Loevinger's or Mokken's coefficients of homogeneity is in my opinion the coefficients most close to latent trait analysis. In our 1992 publication of the aggression scale I asked my statistician, Peter Allerup, to make a latent trait analysis analogue to Mokken's coefficient of homogeneity. Our article illustrated very nicely the parallel thinking in Mokken's coefficient and latent structure analysis.

If I can go back for a moment to our ECT study I would like to give another example of latent structure analysis. In the post-ECT phase we measured the relapse of depression with the Hamilton and the Melancholia scales. In other words, this study gave us the opportunity to measure how depression develops symptomatically. We found that there was the same structure, essentially, as when patients improve during treatment. Hence, anxiety and depressive symptoms came first and then the retardation elements. Guilt and suicide came with those who developed a more severe picture. Cognitive changes came relatively early and introversion or lack of contact was one of the first

512 The Psychopharmacologists II

things when the patients started to relapse. Because paroxetine was best to prevent relapse after ECT we now have the hypothesis that if you give an SSRI continually you can control the anxiety and depressive symptoms and if you can do that you will not get the more severe symptoms.

The hierarchical structure of the aggression scales shows that passive aggression is lowest in the hierarchy of symptoms whereas active aggression has the highest place in the hierarchy. Patients with active aggression have shown signs of passive aggression before their outbursts.

Well there is an interesting point here. In a sense we are very focused on disease entities but of course there is no such thing as aggressive disease. Arguably, however, what we use the treatments for is for aggression and many of our treatments are more effective at controlling aggression than curing diseases.

You are certainly right if you ask why patients come to hospital – it is either because they are too outward aggressive or in the case of the suicidal person they have too much inward aggression. But this is not a diagnosis.

You are certainly right. Duphar with the eltoprazine drug was not asking for a new disease, but rather a scale measuring the target symptoms of aggression. It was Freyhan who in 1959 stated, that the effectiveness of a neuroleptic drug must be measured in terms of its ability to improve the target symptoms. It does not make much sense to relate drug evaluation to diagnostic entities or other generic variables which ignore individual symptomatology. I was happy to have met Freyhan several times in Switzerland at the symposia Professor Kielholz held.

And if we were to say were in the business of controlling aggression it would be the end of the psychiatric profession. We have to say we are treating diseases.

That is right. We had a problem when we decided what to call the scale because in US English aggression is a positive and hostility is the word. In Danish the word aggression means something negative. In the US aggression is positive – you say you do it very aggressively and so on. So we called it also a social dysfunction scale to indicate that there were some problems with it. But you are certainly right and in our country we use clozapine mostly for those schizophrenic patients coming because of aggressiveness to the hospital and in my mind it is not so effective for schizophrenia. It is more of a sedative and anti-aggressive compound and it gives problems if you want to use it for prevention of schizophrenia. Before my time, when I came to where I work now I learnt that it was the part where clozapine has been given there in the 70s before it became problematic. And it was used then mostly for patients with aggressiveness which is often called today a positive symptom in schizophrenia. People now think that Clozapine is perhaps especially good to negative symptoms but I don't think so.

Psychiatric patients have either too much outward directed aggression or too much inward aggression if we by aggression mean hostility or social dysfunction. In US the word aggression has a positive value, meaning something

energetic or dynamic. Our aggression scale, therefore, had to be called social dysfunction and aggression scale (SDAS). The target symptoms of most of our psychopharmacological treatment is to control outward aggression with neuroleptic drugs and inward aggression with antidepressive drugs. We are still not at the end of the psychiatric hospital profession, although it is difficult now to treat depressed patients in hospitals, in UK even more difficult than in Scandinavia. I think.

On the rating area, you have also moved into the area of self-rating.

Yes that comes I used self rating initially because I was interested in the experience of time – my subjects completed the Beck Depression Inventory. But then for some years I did not use self-rating approaches because in the studies when we took the plasma levels when you had to give you a lumbar puncture patients were never motivated to complete a self-rating questionnaire. For the drug studies the nursing staff were not so interested so we stopped using such ratings and we didn't use them for the DUAG studies.

Then, it must have been in 1986, people said to me you must know something about quality of life. I said really I don't know but I will try to look into it. When I went into it, I could see that Goldberg's General Health Questionnaire, Becks Depression Inventory and the Zung depression scales were used in many places as if they were a well-being scale. So it seemed that in the idea of quality of life was on one hand the idea of taking the patient's opinion and the other part of it was to make sure that you had taken all things into account and not only some dimensions but also to try to take into account the goal of treatment. One of the interesting things for me was that we had thought about improvement but quality of life then came to say but what is the goal of treatment – you must also know in the future when to stop treatment and where to go.

If you go into this area, you relatively quickly come to the problem should you make a scale for the patient himself – his own scale. I have tried to look at that. If quality of life is a unique situation for the individual patient, why not then try to do as you do when you start developing a scale – you have some components and you say how should I measure the severity of those components. So we did a study with a beta-blocking agent, which Ciba-Geigy thought had an anti-anxiety effect. It was a study on generalised anxiety disorder compared with placebo and flupenthixol in a small dose. So we had a psychologist who wanted to use a repertory grid technique but Ciba-Geigy were not keen because the psychologist needed so much time at baseline and they thought it would interfere with the drug treatment. So it was decided that it was a six week study and that after six weeks or when they dropped out the psychologist should make the repertory grid and some of the elements would be 'how were you before treatment', 'how are you today' and 'how do you wish to be'. Lundbeck had recommended 2mgs of flupenthixol, whereas in Denmark we usually use 0.5 to 1mg, which may be too much I think. Anyway the result was that on the Hamilton Anxiety Scale flupenthixol was the best for

514 The Psychopharmacologists II

symptom reduction but no patient in this group had a better quality of life. In the placebo group which had the worst effect on anxiety symptoms 30% had a better quality of life. And in the beta blocking group, which was a little bit worse on the symptomatic side, 60% had a better quality of life on this repertory grid. This was not captured by the the Hospital Anxiety Depression scale or the GHQ which we also used.

We were impressed with this and since then we have tried to make a computer version which had taken a long time. I became interested in this idea of going a little bit closer to the patient. There are useful things we can do with a computer model. In the repertory grid, you have components made from the same questions we always ask and patients can add their own important things and from this you do a factor analysis on each patient. Now if you have a computer you can press a button and in 10 seconds you have the factor analysis and a factor score. It has taken us two years to get this working. It was too comprehensive at first and had to be trimmed down because today you must not take more than half an hour – if it takes more than that nobody will use it. Even there we cannot use this family doctors because they want something that takes 5 minutes.

Then of course if you do traditional studies, there are scales such as the SF-36 which of course can be used in all disorders so you can compare the impact of diabetes, epilepsy, depression and schizophrenia. In the US apparently this has had an impact on politicians who can see how much care untreated depressed patients need because all of the components of the scale are decreased in depression. In myocardial infarction of course their somatic daily activities are affected but other things are not so impaired whereas in mental illness everything is.

I became a WHO collaborating centre last year to try to develop outcome measures in the daily clinic. So just now I'm looking at what have I learnt over the years in so-called scientific trials, what use could some of those measurements be in daily outcomes to show how our patients were when they came to the hospital and when they were discharged back to the family doctor six months later.

Yes. I used self ratings initially because I was interested in the experience of time in depressed patients when they completed the Beck Depression Scale. Our data showed that when patients were carefully informed they were able to use this scale. My psychometric analysis focused on a subscale rather similar to that Aaron Beck has published. However, the full scale is most often used. In our plasma level effect trials with imipramine and clomipramine I realised that when the Beck scale was handed to the patient by the nurses without any information the scale was more often empty than completed. I think Malcolm Lader has said that questionnaires can be intrusive to the clinical process, distracting for the patient and orthogonal to clinical impressions. In our DUAG studies we didn't use self-rating scales. It is very interesting that when the clinical trials with the SSRI's were planned in the late 1970's or beginning of the 1980's self-rating scales were often excluded. I have noticed in one of my

papers on meta-analysis of the SSRI's that patients seem to prefer the tricyclics to the SSRI's on the Beck or Zung self-rating scales.

In 1982 I visited Aaron Beck in Philadelphia and followed a one-week course in cognitive therapy. It was the year after my thesis on the Hamilton and Beck scales was published at which Max Hamilton was one of the discussants. During my discussions with Aaron Beck I asked him about the difference of depressive 'symptoms' as measured by his Inventory and 'automatic thoughts' as measured within the negative triad of depression. He gave me no real answer. I still find that the 'negative automatic thoughts' and the 'depressive symptoms' belong to the same construct. When the depressed patient has recovered both the symptoms and the automatic thoughts disappear. Although I received a diploma in cognitive therapy I prefer drug treatment for major depression, both in the acute phase and in the maintenance phase.

Another thing I could not really explain was why Aaron Beck tried to minimize the work of George Kelly on repertory grid. They are both constructivists – they belong to the philosophical school saying that we individually make our own models or constructs of the world because the modern world is too complex for us to experience in toto. Kelly used the term 'constructs' while Beck tried to use the term 'schemas' for these emotional and cognitive models. Thus, 'automatic thoughts' are the depressed person's negative model of his or her world. Beck told me that these construct are cognitions, therefore the term cognitive therapy, although preferred the word emotions.

In 1986 people told me again and again 'you must know how to measure quality of life because you have been working with distress scales like the Beck Depression Inventory, Goldberg's General Health Questionnaire, the Zung scales, etc...' Then I realized that the questionnaires were all measuring the dimension of ill-being versus well-being. At that time quality of life was accepted as a measure of subjective well-being even in such journals as new England Journal of Medicine. Questionnaires completed by the patients themselves were therefore considered an appropriate measure.

I was influenced especially by the study done by Jachuck and coworkers from Newcastle on the effect of hypotensive drugs on quality of life in patients with mild to moderate hypertension. The results of this study showed that while all the treating doctors found that quality of life of their patients had increased during the trial, only half of the patients agreed in the questionnaires, and only in one case the relatives of a patient found an increase in quality of life of the patient.

We tried to replicate this study from Newcastle, but did it only to some extent. During the discussions with patients and their relatives I realized the limitations of standard questionnaires. What was important to one patient might not have any importance for another patient. In this situation George Kelly's repertory grid is very useful. It is essentially a way to develop a scale for each individual patient.

So then we did a study with a beta-blocking agent, which Ciba-Geigy thought had an anti-anxiety effect. It was a study on generalised anxiety disorder.

516 The Psychopharmacologists II

der compared with placebo and flupenthixol in a small dose. We had a psychologist who wanted to use a repertory grid technique but Ciba-Geigy were not keen because the psychologist needed so much time at baseline and they thought it would interfere with the drug treatment. So it was decided that after six weeks or when they dropped out the psychologist should make the repertory grid and some of the elements would be 'how were you before treatment', 'how are you today' and 'how do you wish to be'.

Lundbeck had recommended 2mgs of flupenthixol, which may have been too much – in Denmark we usually use 0.5 to 1mg. Anyway the result was that on the Hamilton Anxiety Scale flupenthixol was the best for symptom reduction but no patient in this group had a better quality of life. In the placebo group which had the worst effect on anxiety symptoms 30% had a better quality of life. And in the beta blocking group, which was a little bit worse on the symptomatic side, 60% had a better quality of life on this repertory grid. This was not captured by the the Hospital Anxiety Depression scale or the GHQ which we also used.

We were impressed with this and since then we have tried to make a computer version because it takes a long time with pen and paper. There are useful things we can do with a computer model. In the repertory grid, you have components made from the same questions we always ask and patients can add their own important things and from this you do a factor analysis on each patient. Now if you have a computer you can press a button and in 10 seconds you have the factor analysis and a factor score. It has taken us two years to get this working. It was too comprehensive at first and had to be trimmed down because today you must not take more than half an hour – if it takes more than that nobody will use it. Even there we cannot use this in general practice because family doctors want something that takes 5 minutes.

Can we do outcomes of research of that kind? If you go back to 1956 when the drugs were introduced and the question of rating scales first came up, Nate Kline said that all these rating scales are like the rabbit in the hat trick. You put the rabbit in the hat and you pull it out and everybody is impressed but in actual fact because of the particular rating scale that you have used you have really created the outcome that you want. What he was saying is that the only outcomes that count are if people actually leave hospital or if they move from the back wards to the front wards. What do you think?

Yes but of course such data you already have – you have the lengths of stay in hospital. But I think that with this SF36 where you compare diabetes, myocardial infarction and other disorders and so on, it would be interesting in hospital to do the same. I have tried to do this and in fact the Department of Health here in the UK has helped develop something similar with HoNOS which is a very simple outcome measurement, which could be used more routinely. I am doing something similar in collaboration with WHO in Europe who have their headquarters and their quality of care department in Copenhagen. In schizophrenia we are looking at people in the hospital and then in the outpatient setting and measuring outcomes using patient satisfaction in some way. It needs to be a small scale with 6 or 7 items from DSM. I think you need sim-

ple scales – you can say perhaps too simple but these should be indicators and if you need more you have to then use further instruments but what we need to do is to go into the daily living with some of these tools.

First of all, when the two most educated psychiatrists, who performed the global severity ratings in my very first study on the Hamilton scale, had finished the study they told me that these ratings often correlated very closely to the patients move from the back wards to the front wards. Now, twenty years later, we have in our country no back versus front wards. The few we have left are mixed! The Hamilton and global ratings are in my view very realistic 'bed-side' instruments.

Secondly, I would like to emphasize that Nate Kline, among the MAO-I discoverers, like Roland Kuhn who was behind the discovery of imipramine had the feeling that the discovery of the antidepressants was purely clinical, without controlled trials, without placebos, without statistics. However, Kuhn has admitted, that he tried in the years from 1956 to 1962 to convince other psychiatrists of the antidepressant effect of imipramine, but without success. It was with the introduction of the Hamilton scale that people first became convinced of the outcome. Max Hamilton has stated that rating scales are not really suitable for exploring a new field of knowledge. Their construction requires much practical experience and an appropriate body of theory – in a sense they are an end-product.

Is what you are saying in one sense that it is one thing to use rating scales to map how people get well and that is what we have been doing but in actual fact it is probably more important to map how or why they are not getting well – turn the problem around in the opposite direction.

That's right, because in the daily routine those who do not get well will disappear but they will turn up in some other place. It is time, I think, to use the knowledge we have gained outside the so called randomised clinical trial.

Can I switch to another trial that you were involved in where I would be interested to hear your views because what we have been talking about in the area of quality of life and in the area of aggression or social dysfunction is very different it seems to me to the alprazolam studies of panic disorder which were very focused on a disease entity – they were very Kraepelinian. You were one of the centres for that and there was afterwards all the controversy that blew up between Klerman and Marks.... How do you read all of that?

Well, first of all I received data from the whole sample in Copenhagen and I was asked to look at the SCL90 because in the first draft of the outcome of this study it was said that it correlated relatively well with the Hamilton scale so they didn't need to go into it. I said I would be very happy to look at it because I always liked the scale. I have actually been in contact with Frank in Baltimore who with Parloff created it. Actually, professor Frank gave me permission to publish the original 41-item version in my book on rating scales. It was actually the first scale to measure the outcome of psychotherapy in the 50s in Baltimore. They said if we cannot demonstrate and document an effect of psychotherapy we have to stop doing it...

518 The Psychopharmacologists II

In the 50s?

Yes, it was a very very nice study – the first study of the SCL90 by Parloff and Frank. They said perhaps we should find something else better than symptoms but they could not find any better language to communicate with anxiety disorder patients so they called it a discomfort scale – which today we call well-being scales. Parloff afterwards became the leader of group therapy at the National Institute. Kelman became Professor in Social Psychology and Frank is the Emeritus Professor in Baltimore. I published the original scale in a book because you can never find it in the literature, so I contacted those people. Originally it was a 30 item scale but their successors had added items up to 90 – most scales started with too many items and work down but this one went up to 90. It is essentially a scale for all kind of psychopathology. My interest was to look for factors because there has been a lot of criticism about the SCL90 and how many factors there are in it – perhaps there is only one or are there 8 or 9 or 5 factors.

So I looked at that part of the study and avoiding the politics, I took only patients on placebo or imipramine and what I found was that the first factor, was a discomfort factor, there was no anxiety factor. The next was phobia and panic came as the third. One of my conclusions was that this must be some quality of life disorder.

Yes, it was a very bright study, the first study of SCL-41 (later becoming SCL-90) by Parloff, Kelman, and Frank. They looked for something else better than symptoms as outcome measures of psychotherapy but they could not find any better language to communicate emotions in their anxiety disorder patients, so they called it a discomfort symptom scale. Parloff became later the leader of group therapy at the National Institute. Kelman became professor in Social Ethics at Harvard and he has published on compliance and the 'crimes of obedience'. Frank is now Emeritus professor in Baltimore and is perhaps best known for his best-seller 'Persuasion and Healing'.

In the cross-national panic study the factor analysis of the SCL-90 identified the original discomfort factor as the first, then came the phobia factor, and as the third factor the panic factor. My conclusion was that the discomfort factor should be considered most important when comparing imipramine, alprazolam and placebo.

Discomfort means what?

Discomfort means, I think, why do people come to a psychotherapist in Baltimore in 1952 – they called it demoralisation also. It was persons who in their daily life felt that had ended up in a corner and they were guilty about it, symptomatically it was depression-like or a lack of well-being which they called discomfort. I sent a paper to a journal with a British referee I think who said that distress was a better word but it was interesting for me at least that distress was the first factor and then came phobia and then came panic.

One of the interesting issues came up in connection with the question of reimbursement for panic disorders. This was discussed in meetings after the results were open, is a disorder in which you have an attack lasting 5-10 min-

utes coming once a week at 8 o'clock on Monday evening perhaps and then disappearing, is that a disease or is it a small problem? Of course you could say given the distress factor that there was something else between the attacks and then there was also the phobic element, so I think it depends very much how you look at the outcome measurements which of the factors of the scales you use. We showed that the effect of imipramine, depending on which factor you took, there was no effect on phobia and so on.

Discomfort was changed to demoralisation when Frank published his first edition of 'Persuasion and Healing'. In this book he defines demoralisation as the state of candidates for psychotherapy, whatever their diagnostic label. It is essentially a coping with illness behaviour, very much analogue to neuroticism according to Eysenck. The discomfort factor had no single anxiety item, but several depression items. I often equate neuroticism and dysthymia, which are indicators of quality of life. Most studies do not find anxiety as an indicator of quality of life. Personally, I always look for depressive symptoms in patients with anxiety disorder.

Klerman more or less said that he was in the business of using the panic disorder studies to internationalise DSM-III and the Kraepelinian view. Is that how you read it at the time? He saw it as a means to spread the use of DSM-III because you would all have to get together and you would use DSM-III criteria.

Well in Europe or in Denmark at the time we did not have a diagnosis of panic disorder so alprazolam could not be approved for this and it was only two years ago when ICD10 had it that it was accepted in Denmark for the indication of panic disorder. But I don't think that this was a good way to show that panic disorders are a Kraepelinian disorder because then you should argue that alprazolam should uniquely treat panic disorders and...

Klerman was not very interested in the SCL-90 findings, he was more interested in the DSM-III criteria of panic disorder. In my own analysis I excluded the alprazolam arm and focused on imipramine versus placebo differences. I found that Klerman focused on alprazolam as a unique drug for panic disorder. However, as a chronic disorder I prefer antidepressants for panic. At the bottom it is the Kraepelinian, medical approach.

It seems to be that he was perhaps not too concerned about showing that alprazolam was the only treatment for this condition but he wanted to sell the condition and with the condition the idea of DSM-III and with that a Kraepelinian view. Perhaps he wanted to sell it mostly back home to the Americans. How do you read the articles in the British Journal of Psychiatry which took 4 years to publish and with them was this amazing correspondence.

To be honest it was too political for me. Scientifically I could not see the problem actually, to be honest. It should also be emphasised that Freud and not Kraepelin described panic disorder a century ago, when he separated anxiety disorders from the American concept of neurasthenia. But I agree with Klerman that affective disorders are biological disorders. The neurologists have been very successful in grabbing Alzheimer's dementia and Parkinson's

520 The Psychopharmacologists II

disease. We should, perhaps, consider hanging onto Kraepelin's depression and Bleuler's schizophrenia.

You were the original mover behind ECNP. How did that come about?

It happened that I was secretary for the Scandinavian Society of Psychopharmacology – I had some years before replaced the Danish pharmacologist Erik Jacobsen, the father of disulfiram, in the committee. When and it had its 25th Anniversary in 1984, I thought that one way to celebrate it could be that during our annual meeting we might have a symposium where I invited the chairman for the different European countries to tell us how psychopharmacology was going in their country. So I invited I think Michael Trimble and Max Hamilton from the UK and Mendlewicz from Belgium, Professor Hippus from Germany, Ballus from Spain, Gottfries from Sweden, Gastpar from Switzerland and so on. We had a nice symposia and then lunch afterwards when it was discussed whether we in Europe should come more closely together. They liked the setting in Copenhagen – all the annual meetings in Scandinavia always take place in Copenhagen. So I said I have done this for many years and it would be easy for me to arrange a preliminary meeting the next year. This was agreed around the table. We set up a small working group of those who had attended. I called the companies and said that we are in a situation where we want to have a meeting based in Copenhagen about a European Society in Psychopharmacology. They asked me 'why, there are already so many associations'. I said that from my own point of view I see that perhaps we could harmonise standards in clinical trials in Europe so that we are not only looking for what the FDA says in the US. They thought this was a good idea and actually I relatively easily raised moneysmonies so I could invite 400 people and pay their cost for coming to Copenhagen, which is where we had the first meeting. Max Hamilton was one of the most active at this. He was very clear about how to set up a new organisation – he had all the rules and things for that. The first official meeting was in Brussels in 1987 but I think we say that the first meeting really was in Copenhagen in May 85.

You say that one of the things that you were interested in was to get some standardisation of clinical trials methods in Europe. There appears to have been something about the early 80s because the AEP and the European Behavioural Pharmacological Society also began then. Was there something about the early 80s that led to the formation of European institutions?

AEP was founded in the same year, although I was told yesterday actually by Professor Pichot that he organised the WPA meeting in Vienna in 83 and after that he discussed with a German psychiatrist the idea of a European organisation, so they started in 84 and I was invited to their meeting in 85. There was some tension – would we be asking the same companies for money, so it was decided that we should have our meetings in ECNP 85, 87, 89 and the AEP the other years. Later ECNP changed to yearly.

AEP was founded at least without my knowledge at the same year as ECNP. However, Professor Pichot has recently told me that the very first initiative was

made in 1983 when he organized with Peter Berner the WPA congress in Vienna. There was some tension when it was clear that both AEP and ECNP was well-established associations, especially concerning sponsorship from the drug companies. It was decided that ECNP should organize their congress every uneven years (1985, 87, 89 etc) and AEP every even year. However, ECNP has later changed to yearly meetings.

Then, I think it was in 89, I asked the committee why shouldn't we go back to our original idea about standardisation because in my opinion we now had a more political society and were looking for nice places to organise the next congress in a more ordinary or traditional way but why couldn't we also focus on other things. There was a discussion in the general executive committee. I at the time had a small sub-committee drawn up on a Scandinavian model – in the Scandinavian Society of Psychopharmacology, we have a sub-group called UKU where we have made a side-effects scale; it is a committee for clinical trials and they receive 10% of the income for making their investigations. So I said couldn't I receive some percentage of our money and create this kind of role in Europe but the executive committee said that I should try to raise the money myself. I said I would do it but it was the same companies I would approach. Anyway I organised a meeting in 1990 in Strasbourg, which was very successful, which was for persons working in the industry in the clinical trial area, trying to set up a dialogue so that they could exchange their knowledge and discuss what problems there were having. It was a very small closed meeting and it later became an independent group because they wanted it to be so. So we still have these small meetings but they aren't part of ECNP.

So did ECNP evolve in a different kind of way to what you had expected or hoped?

One could say that, yes. I had hoped that we could have had more of an influence on giving guidelines etc. Perhaps it was naive from my point of view because the whole thing is of course governed by FDA and in the end things have to be approved in US and the other big countries.

But surely if there were a sufficiently strong European voice on clinical trial methods the FDA will pay heed.

This was what I hoped. At the first meeting in Strasbourg one of the main subjects was quality of life in clinical trials and we had invited representatives from the FDA precisely to be sure that we had a dialogue in the hope that a strong organisation in Europe could be more influential. But ECNP seemed to develop as has ended up as another society for congresses....

I thought the programme at the ECNP meeting in Venice in 1995 was very surprising. There was very little either clinical science or basic sciences – it was all focused on disease entities like Social Phobia or OCD or Panic Disorder. This AEP meeting in contrast shows I would have thought a much broader base.

I left ECNP democratically after 8 years and went actively into AEP. From 1993 to 1994 I was president of AEP and organized in September 1994 the AEP

522 The Psychopharmacologists II

Congress in Copenhagen. I think it was the breakthrough of AEP. This year's joint meeting held with the Royal College in London has been a boost and with Norman Sartorius being the president after Angst I think all this is good.

Related to the question of what happened the ECNP program last year I suspect is something that came up at a recent CINP Members meeting in Melbourne where the debate got very acrimonious. It was said from the floor – and I think most people there probably felt the speaker from the floor had ECNP in mind – that CINP and several of the larger psychopharmacology associations appeared to be run like some of the old European Royal Houses once were; they seem to be in the control of a few people and you wondered to what end they are being run. Does that ring true?

Yes certainly I think that is right as I see it myself.

You mentioned the UKU side-effects scale – how did the UKU side-effects scale actually come about.

The Director at that time for the Health Authorities in Sweden could not understand why in their file for reported side-effects there was nearly nothing – it was as if psychotropic drugs when they were used in Sweden had no side-effects. He asked whether it was because doctors don't know what side-effects were or they accepted a lot of side-effects because that was the daily reality. So UKU consisted of a member from each of the Nordic countries – I was the Danish member – and we were invited by the Health Authority to look at side effects. We decided to make a scale like the Hamilton scale. For instance I invited the Danish expert in sexual dysfunctions to give me some ideas – at that time we could only do that in Denmark. We asked Jens Gerlach for side-effects of neuroleptic drugs etc. Then during our annual meetings we had discussions and then we did a study where we said that all clinics in Nordic countries should use it in their departments for a month for all patients with schizophrenia – we had 2,000 patients. Then we did a small inter-rater reliability study. Later I have also done a small scale for the serotonin syndrome. I am no longer in the UKU for democratic reasons but I think it has been a good thing to do this.

I think it helped to put the whole area on the map.

Yes I still receive letters from people wanting to use it and it has been translated into other languages. It was not perfect but it was a starting point.

One of the unusual things I guess about what you just referred to is that as a group the Scandinavian countries seem to be able to co-ordinate. There is no way that you could tell all the people here in the UK to all use this scale in their clinics. You wouldn't be able to get people to act in concert in that way. How come you guys have been able to?

It was in the years when an older generation of psychiatrists were still in place – Lingjorde in Norway and Dencker in Sweden. The Scandinavian society started in 1959 – it is one of the oldest, the German society may be even older. We used to meet every year. During the period that I was secretary I received

some letters from well established elder psychiatrists in Denmark saying either that they feel now that there is too much industrial influence on the meetings or the meetings have turned too much toward the basic sciences and have become less clinically relevant. I think we did what we did at a time when we still had a lot of clinicians. I think this is declining unfortunately. The members are much more pure pharmacologists; they are not daily acting psychiatrists. But at that time we had access to active psychiatrists around the Nordic countries, who came every year to the meetings and we could report back. So I think today it doesn't work as well.

What were your relationships with Lundbeck like. I know they helped start the Scandinavian Society in 59 and they helped support ECNP....

Well ECNP not so much to my knowledge. It was actually the three Swiss companies in Basle who were most active in ECNP as I remember it. In the Scandinavian Society our UKU meetings were held for many years in the auditorium in Lundbeck. PV Pedersen who was behind most of the Lundbeck drugs was a prominent member from industry. He was a real gentleman. When UKU was originally constructed it was arranged so that the industry could send protocols to be discussed but because there was a Lundbeck influence of course nobody did – why should they have another company etc. So Lundbeck has been very active in the Nordic countries but less so today. The change came about in the mid-80s.

Now I know you are also interested in the concept of hypomelancholia, which has always been an interest of mind – the idea that there might be a milder form of biological/vital depression which perhaps people don't even know that they have. Can you explain to me how you came to the idea? A few different people have written on the same kind of concept but it has never taken off possibly because it is not clear that this is something that needs to be treated.

Yes I am actually giving a lecture today about what I call positive melancholia. We know that of people in society who have a depression, only half of them seek help in the family doctor setting. So we have the Defeat Depression Campaign at the moment and next next year in Denmark we have a Year of the Brain and I am chairman of the depression group and one of our goals is to draw attention to the fact that there are untreated depressed persons. I have speculated why those people do not go to their doctor; is it because it is part of being human that to experience some kind of depression in the long run gives a better quality of life so to speak? You know Kay Jamieson who has published a book on her own experience of manic-depressive illness says that it is so that you have the insanity as the negative things and then some kind of enthusiasm as part of the positive melancholia.

I have looked at the various philosophers from this point of view – although not systematically. I started with David Hume, who in his young years had depression and he actually had letters to his doctor which described the English Malady. From that I looked at William James, T S Elliott, Albert Camus and

524 The Psychopharmacologists II

others who have had depressions, who all say that a depression leads one to speculate about whether life is worth living. So it prompts them to look at their life in a new way and it also gives them some excuse for loneliness and it cannot have been too bad for their creative part to be alone for some periods. William James had what was called the American disease and every time he had this problem he went to Europe for three or six months to England and Germany. So this took him out of the daily routine and enabled him to speculate a little bit about new ideas. Henry James, his brother, in his novels often touches on the hours between 5 and 8 in the afternoon where he said you can know that in those hours you can move from a light depression into perhaps even a small degree of mania. He described those hours as eternity in which he could look into things in a special way. Everyone who is familiar with depression knows that the morning is a little bit worse and that in the evening you come out of it – in this situation you can have a more intense experience of presence, you are more intent to describe things around you

One of the implications of that though is that it might be a mistake to treat this condition.

That's right. My conclusion is that perhaps the right people go to their doctors and those who don't go shouldn't go. I don't know how exact the calculations are but we are often told that half don't go. Also there is the issue that we don't know what may happen when you start treatment – one of the problems can be causing rapid cycling and I have difficulties with these cases – I don't know what to do actually. Many people in the Copenhagen areas refer patients to me and most of the time it goes okay but in this case nobody likes to treat them and when they come to me I don't know what to do. Even with lithium and I have never seen a good effect of anti-epileptics in these cases. I also use thyroid extracts, T3 and T4, but that is not an answer. I have been trying risperidone lately. Now and then I speculate whether it is all the previous treatments which have triggered the problem.

Another group that is awkward to know what to do about it is the group of recurrent brief depressions described by Angst.

I have not seen many of these myself. In a Danish study on Parkinson's disease and depression we started with citalopram and placebo firstly in which one of my psychiatrists would go around and interview using the Hamilton. When the relatives neurologist had some video tapes and could see how to do it, they wanted to do it themselves and then they produced some video tapes for me. I could see that some of those patients could tell that two days ago they had had a depressive episode but it took only one or two days. The study didn't show any effect of citalopram on this and I know there has been two studies one fluoxetine and one paroxetine where no effect has been shown on primarily recurrent brief depression.

Is that because its a different kind of condition?

Well I had also been interested in depression in schizophrenia whether it is again a mental part of the neuroleptic induced Parkinson's syndrome with

more fluctuating recurrent brief depressive episodes. That is another thing we are trying to do work now in my department because we are moving more into schizophrenia. Recurrent brief depression is not a thing I see very much. It is interesting we had Spitzer in Copenhagen to a meeting and he was a little bit sceptical about the nature of this disorder – its not in DSM-IV.

How has the field changed in the last 30 odd years since you began. Periodically you get the view that psychiatrists are doomed to distinction, they are going to be replaced by clinical neuroscientists and psychologists or a combination of those two.

Especially during the last decade I have experienced that a good psychiatrist should be able to work both as a clinical neuroscientist and as a clinical psychologist. The development of the selective and safer antidepressants and antipsychotics need the neuroscientists knowledge of the mechanism of action of such drugs and this is helping the patients and their family doctors to treat the major mental disorders. At the same time psychiatrists have to know about coping strategies both in the minor and major mental disorders which include stress factors and quality of life, i.e. the work in clinical psychology. Only the psychiatrist is able to cover this holistic approach. Otherwise, we are back to the old dualism between brain (neuroscience) and mind (psychology).

It seems that Danish psychiatry in contrast to other European countries, e.g. France and the Netherlands, is mainly based on biological psychiatry?

Yes, when Rafaelsen in 1974 was offered a new chair as professor in psychiatry he wanted it to be in biological psychiatry like Herman van Praag or Julien Mendlewicz. However, the University of Copenhagen would not accept such title because it was the Danish approach in general that psychiatry was biological and therefore the term biological psychiatry was considered a truism. He then became professor of psychopharmacology.

In Denmark we have had very few psychoanalysts compared to other European countries. The most influential was Vanggaard who never became a professor but who worked in the same department of psychiatry as Rafaelsen. They respected each other. Vanggaard was always aware of the limitation of psychoanalysis and used antidepressants in the treatment of major depression. He was, for instance, the first to introduce in Denmark the combination of MAO-I's and TCA's using isocarboxamide and amitriptyline. In monotherapy with TCA's he preferred clomipramine because of its broad spectrum of efficacy including OCD.

All research-active departments of psychiatry in Denmark have been working in psychopharmacology and related areas. Strömngren, the most famous Danish psychiatrist in this century was, however, not only a psychopharmacologist but also an epidemiologist who established the Danish Central Psychiatric Register which now is chaired very successfully by Povl Munk-Jørgensen. He also initiated the national and international family, twin and adoption studies which so successfully have been continued by Fini Schulsinger and Aksel Bertelsen. We have already mentioned Rafaelsen and the generation of clinical psychiatrists he inspired. In this context I should also mention Annette Gjerris.

526 The Psychopharmacologists II

Lars Gram, who is professor of clinical pharmacology is the father of pharmacokinetics and has inspired the next generation, most eminently Kim Brøsen. I should also mention the psychopharmacological research group at Sct. Hans Hospital in Roskilde. The first generation included Faurby and Munkvad. The second generation includes such names as Rasmus Fog who has looked at the role of dopamine in schizophrenia and Jes Gerlach who especially works with tardive dyskinesia. The idea that psychiatry basically is biological has guided clinical research in Denmark in this century. It may explain the scientifically relatively high level and standard of Danish psychiatrists.

References

- Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG (1975). Quantitative rating of depressive states. *Acta Psychiatrica Scand* 51, 42-50.
- Bech P (1975). Depression: Influence on time estimation and time experience. *Acta Psychiatr Scand* 51, 42-50
- Bech P, Rafaelsen OJ, Kramp P, Bolwig TG (1978). The mania rating scale: scale construction and inter-observer agreement. *Neuropharmacology* 17, 430-431
- Jachuck SJ, Brierly H, Jachuck S, Willcox PM (1982). The effect of hypotensive drugs on the quality of life. *J Roy Coll Gen Pract* 32, 103-105
- Scandinavian Association of Psychopharmacology 1959-1984. *Nord Psykiatr. Tidsskr Supplement* 10. (1984) vol 38.
- Bech P (1985). European College of Neuropsychopharmacology. *Pharmacopsychiatry* 19, 1.
- Danish University Antidepressant Group (1986). Citalopram: clinical effect profile in comparison with clomipramine. A controlled multicentre study. *Psychopharmacology* 90, 131-138.
- Lingjaerde O, Ahlfors UG, Bech P, Dencker SJ, Elgen K (1987). The UKU side effect rating scale. *Acta Psychiatrica Scand* 76, supplement 334.
- DUAG (1990). Paroxetine. A selective serotonin reuptake inhibitor showing better tolerance but weaker antidepressant effect than clomipramine in a controlled multicenter study. *J Affective Disorders* 18, 289-299.
- Frimodt-Møller J, Loldrup Poulsen D, Kornerup HJ, Bech P (1991). Quality of life, side effects and efficacy of lisinopril compared with metoprolol in patients with mild to moderate essential hypertension. *J Hum Hypertension* 5, 215-221
- European Rating of Aggression Group (ERAG) (1992). Social dysfunction and aggression scale (SDAS-21) in generalized aggression and in aggressive attacks: A validity and reliability study. *Int J of Methods in Psychiatric Research* 2, 15-29.
- Lauritzen L, Clemmensen L, Klysner R, Loldrup D, Lunde M, Schaumburg E, Waerst S, Bech P (1992) Combined treatment with imipramine and mianserin. A controlled pilot study. *Pharmacopsychiatry* 4, 182-186
- Bech P, Allerup P, Maier W, Albus M, Lavori P, Ayuso JL (1992). The Hamilton Scales and the Hopkins Symptom Checklist (SCL-90). A cross-national validity study in patients with panic disorders. *Brit J Psychiat* 160, 206-211
- Thunedborg K, Allerup P, Bech P, Joyce CRB (1993). Development of the Repertory Grid for measurement of individual quality of life in clinical trials. *Int J of Methods in Psychiatr Res* 3, 48-58
- Bech P (1993). Rating scales for psychopathology, health status, and quality of life. A compendium in accordance to DSM-III-R and WHO systems. Berlin, Springer

Measurement & organisation in psychopharmacology 527

- Bech P (1996). The Bech, Hamilton and Zung Scales for mood disorders. Second Edition. Berlin, Springer
- Lauritzen L, Odgaard K, Clemmesen L, Lunde M, Öström J, Black C, Bech P (1996). Relapse prevention by means of paroxetine in ECT-treated patients with major depression: A comparison with imipramine and placebo in medium-term continuation therapy. *Acta Psychiatr Scand* 94, 241-251