Thomas A. Ban: Neuropsychopharmacology in Historical Perspective

Collated 35

Lithium
4. Controversy
Barry Blackwell vs Paul Grof and Jules Angst

Contents

Barry Blackwell: The Lithium Controversy: A Historical Autopsy

I am delighted Larry Stein has joined Jose de Leon in expressing interest and concern about aspects of an ancient controversy that may have contemporary relevance. Perhaps it is time to engage in a more detailed and complete analysis of the issues raised, many of which are dealt with in my memoir, *Bits and Pieces of a Psychiatrist's Life,* and will be cited in this essay (Blackwell 2012).

It is now almost half a century since Michael Shepherd and I published our article "Prophylactic Lithium*;* Another Therapeutic Myth?" in the *Lancet*, which commented on and critiqued a previously published study by Mogens Schou and his colleague in the Archives of

General Psychiatry (Baastrup and Schou 1967), making the claim that lithium had a unique effect in preventing future episodes of manic depressive disorder. Their riposte to our critique appeared later the following year (Baastrup and Schou 1968).

If history has anything to offer today, then such past events deserve to be dissected. As possibly the sole remaining protagonist in the fierce debate these two papers generated, I offer this autopsy, personally performed and invite INHN members to comment.

This essay will be in three parts: reciting the facts themselves; an analysis and interpretation of the scientific zeitgeist prevailing at the time, commenting on the emotions aroused; and, finally, the possible relevance of such matters today.

I completed five years of psychiatric training at the London University Institute of Psychiatry and Maudsley Hospital, including a two-year fellowship in animal research leading to my doctoral degree in Pharmacology from Cambridge University. Following this, I completed a two-year research fellowship with Michael Shepherd. At his suggestion, I undertook to analyze and critique Schou's data claiming that continuous administration of lithium prevented future episodes of manic depression. There was no control substance since other "mood stabilizers" were far in the future and Schou rejected placebo as unethical based on his clinical experience and convictions of efficacy. So, there was no double-blind procedure to protect against potential observer bias, although a placebo control was included in the definitive studies that confirmed his beliefs many years in the future. The possibility of bias existed both due to the study design and because Schou was quite open to admitting enthusiasm for his hypothesis, derived from a family member's benefit after all else had failed to stifle recurrences. At this time, prophylaxis was such a unique and unexpected claim it might have evoked a "too good to be true" skepticism, which heightened our concern about potential bias in an uncontrolled study.

There was no established method, at this time, with which to evaluate such a unique claim; Schou's series included a heterogeneous collection of subjects broadly interpreted as suffering from manic depressive disorders but with varying affective manifestations, of differing duration, frequency and severity. This created concerns about the specificity of the claim as well as statistical issues, primarily concerned with regression to the mean – spontaneous remission from a high baseline in a fluctuating disorder. Other statistical concerns were displayed and discussed in sophisticated terms in a paper read to a NIMH/VA study group and subsequently published in Frank Ayd's newsletter (Blackwell 1969). Similar statistical and methodological criticisms were

made by Malcolm Lader in the *Lancet* (1968). The essence of these concerns focused on the impossibility of distinguishing dependency on a medication, or spontaneous remission from prophylaxis, a problem I dubbed the "panacea paradigm." The scientific caveats evoked sharp rebuttals from clinicians who knew better, including Nate Kline in America (Kline 1968) and Sargent in Britain (Sargant 1968). Sargant's comments are especially illustrative of the tone and angst aroused in this debate. He appealed for the abandonment of "crude statistics" and "valueless double-blind sampling" in favor of "bedside observations for the sake of England's treatment reputation in world psychiatry."

Seldom noted or commented on is that in addition to concerns about methodology we applied Schou's statistical technique to a convenience sample of 13 manic-depressive patients from the Maudsley data base treated with imipramine and found results comparable to lithium.

It is important to place these events in their broader historical perspective and consider how this colored the controversy. Until the Flexner revolution in the early 20th century, medicine was an apprentice profession whose *materia medica* included many panaceas, nostrums and placebos, the popularity of which depended largely on the status of the apothecaries, physicians or barber surgeons who dispensed and endorsed them. As medicine became more scientific and moved from the community into academic medical centers, its remedies became potentially more effective. Trial methodology and statistical analyses developed to rigorously evaluate therapeutic claims. Eventually, the double-blind controlled study became the gold standard. Psychiatry lagged behind in this regard; chloral hydrate, barbiturates, paraldehyde and amphetamines were synthesized and well established with regard to effectiveness and shortcomings but nothing new or potentially more effective existed to compare them against.

Lithium had a persisting role in this evolution. A naturally occurring metallic ion with no commercial potential or synthetic rivals, it was introduced into medical practice in 1859 as a *bone fide* treatment for gout but then increasingly as a panacea with Lithia tablets used for a wide variety of ailments, despite absence of benefit and occurrence of side effects. In the earlier days of scientific medicine, it was used as a salt substitute in cardiac disease until the absence of a method for measuring blood levels led to cases of fatal toxicity. It was withdrawn from medical practice in 1949, the same year Cade reported its therapeutic effect in psychotic manic patients.

Many pioneers in psychopharmacology consider the two decades from 1950 to 1970 as the seedbed for all the original treatments in every category of psychiatric disorder. Lithium provides

twin bookends for this exciting epoch, beginning with Cade's discovery of lithium for acute mania and ending with Schou's discovery of prophylaxis - both enabled by discovery of a method for measuring lithium levels in the blood. In an account of his own discovery, Cade recognizes Schou as "The person who has done most to achieve this recognition."

The trajectory of lithium's ascendancy as a prophylactic agent during these two decades is best told by Schou himself (Schou 1998) and Paul Grof, with whom he collaborated (Grof 1998) and who wrote Schou's obituary at the time of his death in 2005 at age 87 (Grof 2006). The obituary is an appropriate paean of praise for a colleague who was twice nominated for the Nobel Prize in medicine and physiology. Grof traces Schou's dedication to our field from vivid childhood memories of depressed patients in the asylum where his father was medical director, "wandering in the hospital park with drooping heads and melancholic faces waiting for the depression to pass and fearing future recurrences." This impressed on Mogens the need for a sustained prevention of depression "at the time when maintenance ECT was clearly not the ideal."

When Cade published his findings on lithium in 1949, it attracted Schou's attention although Cade himself had only demonstrated an acute effect in manic psychosis and found that "in three chronically depressed patients, lithium produced neither aggravation nor alleviation of their symptoms" (Cade 1971). Despite this fact, Schou's interest was piqued by his concern that since age 25, his brother had experienced "yearly episodes of depression. In spite of ECT, drug treatment and hospitalization the depressive attacks came again and again" (Schou 1998). During the decade 1950-1960 that Cade vigorously pursued his interest and research on lithium, imipramine was probably not available until towards the end of the decade and it is likely that during this interlude, Schou prescribed his brother lithium, which "changed his life and the lives of his wife and children." This leads me to wonder if, in fact, his brother manifested a Type 2 bipolar disorder, in which mild hypomania went unremarked. Grof notes that late in his career, Schou developed a special interest in "hidden bipolars" – patients with depression who had unrecognized bipolar disorders. Schou's last scientific presentation, shortly before his death, was on this topic and a new study he was proposing (Grof 2006).

Schou was not a founding member of the CINP but participated in the first Congress in Rome in 1958 when he contributed to the final session a "General Discussion." He recalls his comment that "On the chemotherapeutic firmament lithium is one of the smaller stars" (Schou 1998). Baastrup and Schou's seminal publication in the *Lancet* (Baastrup and Schou 1968) had

been underway for seven years, begun probably in 1961. The above facts help explain why imipramine was not included as a comparative drug, even though the population included both unipolar and bipolar depressed patients. Later on, as his familiarity with imipramine grew, he used the term "normothymics" to include both lithium and imipramine (Schou 1963).

These events resonate with the concerns raised in our paper criticizing Baastrup and Schou's methodology and conclusions (Blackwell and Shepherd 1968) regarding the uncertain specificity of lithium and the absence of a control comparison. To be fair, Schou and Grof draw attention to the problem of using a placebo control based on the high suicide rate in untreated affective disorder. Schou eventually resolved this obstacle with a novel trial design in which sequential analysis of paired placebo and lithium patients was coupled with an immediate switch to open treatment for any recurrence(Schou 1998).

Because the *ad hominem* aspects of this debate still linger, I will quote a few laudatory comments made by his friend and colleague Paul Grof in the obituary. Schou was "a caring man with great humility," with a "love and compassion for people" and also a "highly meticulous" researcher who "never left a task undone."

In 1970, two years after I immigrated to America, my mentor Frank Ayd and I conceived the idea to invite all the scientists and clinicians who had discovered the original therapeutic compounds in each disorder to tell their own story at a conference in Baltimore. These first-person accounts were published the following year in our edited book, *Discoveries in Biological Psychiatry* (Ayd and Blackwell 1971). They included Albert Hoffman *(Hallucinogens*), Frank Berger *(Meprobamate),* Irv Cohen *(Benzodiazepines),* Pierre Deniker *(Neuroleptics),* Nate Kline (*MAO Inhibitors),* Roland Kuhn *(Imipramine),* John Cade *(Lithium),* Paul Janssen *(butyrophenones)* and Jorgen Ravn *(Thioxanthines).* I contributed a chapter on "The Process of Discovery" using the interaction of cheese and the MAOI as a template and Frank Ayd concluded with a summary on "The Impact of Biological Psychiatry."

Noteworthy now, but not discussed at the time, was that Frank did not include Schou. Perhaps, speculatively, this might have been for two reasons. First, Schou's contribution was derivative to Cade's and more adaptive than original; secondly, because the benefits of all these "serendipitous" discoveries had all been confirmed in well controlled clinical studies. The methodological difficulty of proving prophylaxis and the specificity of lithium in doing so, would linger experimentally (but not in practice) for almost 20 years, until the definitive studies, in 1984,

by the Medical Research Council in Britain (Glen, Johnson and Shepherd 1984) and the NIMH study group in the USA (Prien, Kupfer, Mansky et al. 1984). This latter study, larger of the two, involved a two-year follow up of 117 bipolar and 150 unipolar patients given lithium, imipramine, both drugs or placebo. It reached three major conclusions:

1. Imipramine is preferable to lithium for long term prevention following recovery from an acute episode of unipolar depression.
2. For both bipolar and unipolar disorders, the preventative effects of both lithium and imipramine parallel their effects in acute episodes.
3. Even when lithium and imipramine are effective they are not panaceas. Only a quarter to a third of patients with either bipolar or unipolar disease were treatment successes.

Eighteen years after Schou's original study, the issues of diagnostic specificity, comparative and specific benefits for lithium or imipramine and their magnitude were scientifically defined in the absence of potential observer bias and statistical flaws.

In retrospect, some of the angst directed toward Shepherd and I might have emanated from various attributions: methodological puritanism, unjust allegations of bias or of potential therapeutic nihilism - for which the Maudsley was rather unjustly credited. Nevertheless, it was a contemporary and colleague of mine from the Maudsley who, in comments on events in the 1960s, made the satirical observation that, "Writing from the Olympian heights of the Institute of Psychiatry Barry Blackwell and Michael Shepherd airily dismissed Schou's evidence" (Silverstone 1998). But we were all scientific babes in the wood when it came to prophylaxis, bias must always be assumed unless it is eliminated and, while the atmosphere at the Institute was decidedly empirical, it was also benevolent to developments in psychopharmacology. The 1998 book, *The Rise of Psychopharmacology and the Story of the CINP,* lists the 33 Founders of the organization; 27 were clinicians but only three were from Britain, Sir Aubrey Lewis, Michael Shepherd and Lindford Rees. Sir Aubrey was an active participant in the first CINP Congress.

My first rotation at the Maudsley as a resident, in 1962, was under Lindford Rees, a dedicated psychopharmacologist who carried out early studies on imipramine; my second rotation was on the Professorial Unit, where Aubrey Lewis took me under his wing and, once he was sure I was not interested in psychoanalysis, arranged and endorsed my psychopharmacology training. True, Michael Shepherd was a skeptic and scientific purist but, lest he be blamed for any perceived

disrespect towards Schou, I must make clear that I was first author on our *Lancet* paper, chose its title and was responsible for the data analysis and conclusions reached.

Nor were either of us wedded uncritically to double blind methodology. We were well aware of its shortcomings. Immediately before our paper on lithium, Shepherd and I worked on a drug study for a pharmaceutical company which went nowhere because of rigid, impractical and unrepresentative criteria for recruiting subjects. We published our conclusions on contemporary trial methodology in the *Lancet* (Blackwell and Shepherd 1967). During my psychopharmacology research in animals, I collaborated with a colleague evaluating and recording the outpatient use of MAO Inhibitors by all the consultants and residents at the Maudsley. This must have been among the first "effectiveness" studies to look beyond the boundaries of conventional controlled clinical trials at what happens in real life (Blackwell and Taylor 1967). The results were unusual and revealing. One intriguing finding was how the interaction between prescriber and drug influenced outcome, precisely what the double-blind study is designed to stifle or eliminate. The most powerful effect on outcome, above diagnostic and demographic variables, was prescriber behavior. Those who used MAOIs a lot, as "first choice" drugs," had better outcomes than those who used them more reluctantly, as "second choice" drugs. The reasons appear self-evident. First choice prescribers reaped the benefits of their enthusiasm, the placebo response, spontaneous remission and perhaps a willingness to tolerate side effects. The "second choice" population contained more treatment resistant and side-effect sensitive patients alert to the physician's skepticism. Needless to say, these outcomes were likely to reinforce physician attitudes and behaviors. Pharmaceutical reps soon learned to capitalize on this phenomenon by offering physicians a stipend in return for using their new drug in "the next few patients you see."

Another finding was the intriguing comment one enthusiastic prescriber made in the chart, "Although this patient never looked depressed before, she looks less depressed now." Perhaps drug outcomes sometimes influence diagnostic habits. So, in retrospect, one wonders if Schou's late-life interest in "hidden bipolars" was evoked by his extensive experience and enthusiasm for lithium. Perhaps he was curious to find if there were subtle and covert clinical indicators of hypomania in some recurrent unipolar patients who, like his brother, unexpectedly benefited from lithium.

Also relevant to the prophylaxis debate was our finding that 18% of that population remained on an MAOI for three years after recovering from an initial episode of "atypical"

depression and relapsing on attempts at withdrawal, a finding we attributed to "dependence" but identical to the 11 out of 60 patients (18%) who took lithium for three years and where "prophylaxis" was the explanation (Baastrup and Schou 1967). Further complexity is added by noting that, independent of diagnosis or treatment method, about 80% of all outpatients at the Maudsley stopped treatment within three months, while the remaining 20% remained, sometimes for years. What then is the difference between "dependency" and "prophylaxis?" This raises semantic, philosophical and clinical issues and attempts to discriminate by stopping treatment introduce an ethical dimension of potential harm. Perhaps this introduces an "eye of the beholder" component concerning which semantic meaning one applies and is this, in turn, partly based on the physician's temperament?

I am ambivalent; my heart tells me one thing and my head another. Am I a neutral researcher, seeker after truth, or a benevolent healer following the Hippocratic ideal of "first do no harm"? Is what I see as "prophylaxis" or "dependence," perhaps some of each?

The issue of potential clinical bias is nuanced; an intimate interaction between clinician and patient, particularly a friend or relative, can sow the seed of a new idea, worthy of further investigation or testing as a hypothesis. The problem arises in how to remove this bias towards the new idea from the outcome of an investigation. Sometimes it is more difficult than others and in my own initiation into research I was fortunate.

As a first-year resident, I became involved in the interaction of MAOI and tyramine containing foods. The first clue to the possible cause of a sometimes-fatal hypertensive crisis came when a hospital pharmacist (GEF Rowe) read a letter I wrote to the *Lancet* describing the syndrome and its symptoms – predominantly a sudden, severe pounding headache. He recognized and described this process in his wife on two consecutive occasions after she ate cheese: "Could there be something in the cheese?" So, a fellow resident and I took an MAOI for two weeks before eating cheese from the hospital cafeteria. Nothing happened. Nevertheless, I subsequently obtained data from 12 cases in less than 9 months, some including measures of blood pressure and one produced under experimental conditions (Blackwell 1963). Nobody suggested my interest and potential bias was artificially elevating a patient's blood pressure or causing a headache. But the research director of the pharmaceutical company making the MAOI did write a letter to the *Lancet* stating that my conclusions were "unscientific and premature." Within weeks, researchers at another hospital had isolated tyramine in their body fluids after eating cheese. The issue was no

longer moot. Physiological and physical parameters are less subject to observer bias than emotional and behavioral outcomes but finding a glib reason to disparage either is easy.

The issue at stake is also a matter of semantics and timing. The word "bias" has a pejorative connotation, especially when applied retrospectively, to allege an investigator's potential faulty judgment in an uncontrolled study. The term then assumes an unpleasant but perhaps unintended *ad hominem* element. Contrast this with the prospective benign intent of a controlled study - to protect an investigator from his or her laudable compassion and therapeutic enthusiasm.

On which side of this semantic fence one sits, at a given moment or on a specific issue, may be influenced by other factors, including the reputation and fame of the investigator and one's acquaintance with them or sympathy with their claims or ideas. There is no better example than Linus Pauling's orthomolecular beliefs and zeal in promulgating them. He was the only scientist to have won two unshared Nobel Prizes: Chemistry, in1954, and the Peace Prize, in 1962. No person on the planet had better scientific and humanistic credentials. But following the onset of Bright's disease, he developed a strong belief that physical and mental illness might be alleviated by manipulating vitamin levels. In 1968, he published an article in Science on "Orthomolecular Psychiatry." Pauling, himself, took three grams of Vitamin C daily to prevent the common cold and collaborated with a British cancer surgeon on its use in prolonging life. These claims were not disproved until more than 10 years later by controlled research at the Mayo Clinic. A physician critic, in an article in *The Atlantic* (Offit 2013) commented that although Pauling was "spectacularly right" in his early scientific career, his late-career orthomolecular assertions were "so spectacularly wrong that he was arguably the world's greatest quack." Putting this cautionary tale aside, it is only just to remark that Schou was certainly right, while Pauling was unequivocally wrong.

By the time Schou was attempting to demonstrate the prophylactic potential of lithium in Scandinavia, the Congress in the United States had enacted the Harris-Kefauver legislation mandating that drug manufacturers prove their products were effective as well as safe. In 1968, I immigrated to America to become the Director of Psychotropic Drug Research for the Merrell Company in Cincinnati. The company was just recovering from the stigma of having marketed thalidomide for insomnia and the marketplace was cluttered with compounds in search of a credible rationale or proof they were more effective than a placebo. Merrell had two such products in the psychotropic domain and I had the daunting task of proving they could pass muster. One

was Alertonic a cunningly named reddish-brown liquid popular in nursing homes for the elderly that contained small amounts of alcohol, B vitamins and an amphetamine-like stimulant. A substantial placebo response made the task of proving efficacy impossible.

A still more dubious drug was Frenquel with the marketing claim that it stifled hallucinations whatever the diagnosis and the odd characteristic that the intravenous dose was higher than the oral one. Since no other drug had a similar claim, this was a niche product and the threat of withdrawal produced a flood of protests from patients and clinicians who "could not live without it." The FDA was unimpressed and impervious to testimonials, but I decided to visit one of the more credible supplicants to better define what was going on. The following account appears in my memoir in the piece on "The Pharmaceutical Industry*"* as a Bit titled "Snake Oil" (Blackwell 2012).

> "I had a trip planned for New York and decided to call on one of the Frenquel seekers. The office where the cab let me off in Greenwich Village was next to a homeless drop-in center. The doorbell was answered by a polite, casually dressed, older physician who greeted me and ushered me into a room in the basement furnished more like a family doctor's office than a psychiatrist's den. In the center of the room stood an examining table rather than a reclining couch with an attached shiny aluminum tray on which lay a large syringe containing a colorless liquid I assumed was Frenquel. Sitting on the table, legs dangling and wearing a brightly colored, mildly revealing dress was an attractive young woman. Almost before I could take in the scene, she leapt to the floor, faced me and began to shout, 'So you're the f----ing drug company man that's going to ruin my life!'
>
> "The doctor moved quickly to take her arm, guided her back to the table, and did his best to calm her. She settled down and lay back, still eyeing me furiously, pulling up the sleeve of her dress to expose the veins in the hollow of her arm. This was obviously a well-practiced routine, which the doctor performed often. He inserted the needle and gently pushed the plunger as the patient closed her eyes and appeared to drift into a light sleep. Visibly relieved the doctor removed the needle, lay down the syringe and leaned towards her.

> 'It's all right, Martha, you can get up now.' Her eyes opened, she smiled at us, and thanked me for coming so far out if my way to help her.
>
> "Another surprise awaited me; the doctor suggested the three of us have lunch together. We walked to a nearby bistro and, over a meal paid for by Merrell, I spent an hour in the company of two friendly, apparently normal people. Over lunch the doctor explained to me that the alcohol and drug detox clinic adjoining the homeless center used Frenquel often to help 'bring down' people in drug withdrawal.
>
> "On the flight back to Cincinnati, I wrote up my "trip report" explaining I had found two 'off-label' novel uses for Frenquel: to calm someone who, most likely, had a borderline personality and to facilitate drug or alcohol withdrawal. I didn't suggest Merrell pursue research into these potential new indications, but perhaps I was wrong. New uses for old drugs are often discovered by chance; looking for one thing and finding another. It's called serendipity. On the other hand, it seemed more likely that everything attributed to Frenquel might be due to suggestion, the placebo response, or spontaneous remission."

I did not state the obvious – that Frenquel clearly had mild sedative and calming properties but certainly not sufficient to justify the rigors of a controlled study in a market already including meprobamate and the first benzodiazepines. Nor were Alertonic and Frenquel a worthy match for lithium in the effort it would take to prove they were effective remedies for a specific problem.

Finally, we come to the saddest part of this tale – the extent to which scientific disagreements can degenerate into strident squabbles. Almost twenty years after our *Lancet* article, Michael Shepherd asked me to review the book, *The History of Lithium Therapy* (Johnson 1984), published in *Psychological Medicine* the following year. The author, an academic psychologist, had authored three previous texts on lithium and claimed Schou and Cade as his friends. In unrestrained hyperbole, verging on the ludicrous, he endorses the enthusiasts who see lithium as "the King of drugs" responsible for the "third revolution in psychiatry." The following quotations illustrate the polemical nature of the book:

- Lithium is being taken by "one person in every two thousand in most civilized countries" because "depression (sic) is a crippling condition."

- Lithium alone triggered the chemical revolution in psychiatry; "At a stroke, the elusive ethereal Freudian psyche was replaced as the primary object of attention in psychiatry by the polyphasic, physic-chemical system called the brain."
- Lithium, "like no other single event, led to psychiatry becoming truly interdisciplinary." Its ubiquitous use "suggests a new basis for classification of psychopathological states." And it is so cheap and easy to administer it will "transform health care in underdeveloped countries."

These absurd claims provoked me to satire and to ending my review by suggesting that those who might buy the book would be those who shared the author's view that lithium was the "Cinderella of psychopharmacology" and who wished to have an unabridged version of the fairy tale at their fingertips. These comments were, in part, a reprise of a lively debate between Nate Kline and me in the correspondence columns of the *American Journal of Psychiatry.*

The final irony is that this book was published shortly before the two definitive controlled studies (referred to previously) finally arrived at an accurate scientific demonstration of the specific and fairly modest benefits of lithium and imipramine in preventing recurrences of bipolar and unipolar disorders, respectively.

Some reservations about the impact of unbridled enthusiasm for prophylactic treatment have been expressed from the scientific sector. Paul Grof notes that the use of prophylactic treatment for "nearly everyone with recurrent affective disorders has led to the point that the natural history of affective disorder the illness is not known anymore. He also notes that with the extensive use of lithium "the concept of affective disorders has dramatically broadened and mood symptoms, rather than comprehensively assessed psychopathology have become the center of psychiatry assessment." (Grof 1998). It is worth adding that the parsimony of the DSM system has colluded in this outcome.

What can we make of all this today? To begin with, the testing of new psychotropic drugs has passed almost entirely out of the hands of academic clinicians and federally funded projects and into the realm of the pharmaceutical industry and subcontracted commercial companies who, while they adhere to FDA minimal requirements for controlled studies, have adopted other dubious ways to degrade the process and bias the outcomes. We have also learned that even the best of controlled double-blind studies may not mirror or predict what happens in real world effectiveness. I would gladly return to the time when experienced dedicated clinicians like Mogens Schou did

the very best they could, however imperfectly, to show us what works in real practice. After all, their original study was really an "effectiveness" one and not a controlled scientific evaluation. And Schou was, after all, correct. But perhaps Mogens Schou's legacy is better served by the recognition that his truly innovative contribution was the concept of "prophylaxis" itself and not the agents used to accomplish it. This was the very fact that relentlessly recurrent episodes of affective disorder could be checked by continuous, rather than episodic treatment, a technique that also suppressed the phenomenon of kindling.

Now we come to the most tantalizing question raised by this autopsy. Suppose that each of us, Schou, Shepherd, Blackwell and Grof are double-blind neuroscientists groping the same elephant. That prophylaxis of recurrent affective disorders is Schou's reality - *the body*, but that lithium is not a panacea for all its forms (Blackwell and Shepherd) - *the tail*, and that more scrupulous analysis of the phenomenology, genetics and neurochemistry might reveal which subtypes respond specifically to lithium, imipramine or valproic acid (Grof) - *the head*. This is a puzzle beyond the capacity of DSM 5 or contemporary trial methodology to solve; worse still, all three compounds are orphan drugs – either un-patentable or generic, so that support for research is unlikely unless the national or federal funding agencies in Britain and America reverse course and revive clinical psychopharmacology research.

At the same time, claims that exceed the level of proof available in efficacy or effectiveness studies should always be challenged and those who exaggerate them beyond belief are free game for Anglo Saxon satire. *Mea culpa!*

References:

Ayd F, Blackwell B, editors. Discoveries in Biological Psychiatry 1971, reprinted Ayd Medical Communications, Baltimore, 1984.

Baastrup PC, Schou M. Lithium as a Prophylactic Agent: Its Effect Against Recurrent Depressions and Manic-Depressive Psychosis. Arch. Gen. Psychiat. 1967;16:162.

Blackwell B, Shepherd M. Early Evaluation of Psychotropic Drugs in Man. Lancet 1967;2:819-22.

Blackwell B, Taylor D. An Operational Evaluation of MAOI. Proceedings Royal Soc. Med., 1967.

Blackwell B. Hypertensive Crisis Due to MAOI. Lancet 1963;2:849-51.

Blackwell B, Shepherd M. Prophylactic Lithium: Another Therapeutic Myth? Lancet 1968;1:968.

Blackwell B. Lithium: Prophylactic or Panacea? Medical Counterpoint: 1969;52-59.

Blackwell B. Bits and Pieces of a Psychiatrist's Life. Xlibris, 2012.

Glen AIM, Johnson AL, Shepherd M. Continuation Therapy with Lithium and Amitriptyline in Unipolar Depressive Illness; a Randomized, Double-Blind Controlled Trial. Psychological Medicine 1984;14:37-50.

Grof P. Fighting the Recurrence of Affective Disorders. In: The Rise of Psychopharmacology and the Story of the ACNP, Ban TA, Healy D, Shorter E, editors. CINP 1998;101-4.

Grof P. Obituary: Mogens Schou (1918-2005). Neuropsychopharmacology 2006;31:891-2.

Johnson FN. The History of Lithium Therapy. The Macmillan Press: London. 1984. (pp. 198).

Lader M. Prophylactic Lithium? Lancet 1968;2:103.

Prien RF, Kupfer DJ, Mansky PA, Small JG, Tuason VB, Voss CB, Johnson WE. Drug Therapy in the Prevention of Recurrences of Unipolar and Bipolar Affective Disorders. Report of the NIMH Collaborative Study Group. Arch. Gen. Psychiat. 1984;41:1096-104.

Sargant W. Prophylactic Lithium? Lancet 1968;2:216.

Schou M. Normothymotics, "Mood Normalizers". Are Lithium and Imipramine Drugs Specific for Affective Disorders? Brit. J. Psychiat. 1963;109:803.

Silverstone T. Psychopharmacology in the 1960's. In: The Rise of Psychopharmacology and The Story of the CINP, Ban TA, Healey D, Shorter E, editors. CINP 1998;285-8.

June 19, 2014

Paul Grof's and Jules Angst's comments

Somewhat Different Hindsights

Paul Grof's comment;

Renaissance of interest

For quite a while, lithium treatment had fallen out of favor in the mainstream. Non-patentable and inexpensive, lithium could not compete with the skillful marketing of new

profitable neuroleptics and antiepileptics and could not withstand other pressures exerted by the pharmaceutical industry. The finest example was the clever advertising of divalproex which, despite the absence of evidence for stabilizing patients, quickly became the best-selling drug for bipolar disorder in the United States. But recently, a renaissance of interest in the use of lithium treatment has unexpectedly emerged.

Several motives may be converging here. Lithium's rather unique antisuicidal properties, proven for some time (Müller-Oerlinghausen, Ahrens, Volk et al. 1991) have recently been widely publicized. In neuroscience laboratories lithium has turned out neuroprotective (Hajek, Cullis, Novak et al. 2012) and it might even become helpful in the management of several obstinate neurological and geriatric disorders (Quiroz, Drevets, Henter and Manji 2012). More important clinically, the re-evaluation of atypical neuroleptics in the treatment of bipolar disorders has lately curved sour. Furthermore, voices have now arisen suggesting that lithium may actually be the only true mood stabilizer, as it demonstrably acts against both polarities of manic-depressive disorder (Grof and Müller-Oerlinghausen 2009).

Perhaps it was this resurgence of interest that led colleagues to ask independently Barry Blackwell and myself to address again the history of the lithium controversy. And, as Barry Blackwell completed his interesting reminiscences (Blackwell 2014) and invites comments on his version of the autopsy, I happily oblige.

It is in this context that I think it is useful to dissect the lithium controversy. I concur with Blackwell that we can learn from the past. In psychiatry we now live in an era of conceptual turmoil and absorbing lessons from our history has become critical. Each story has at least two ways of interpreting. With the passage of time our differences have softened, and I agree with most of what Barry Blackwell says in general but still part with him on the weighing of the usefulness of long-term lithium treatment.

Preceding events

What was the controversy actually about? Let me first briefly sum up, from my perspective, the events that preceded the disagreement. In the 1950s the maintenance treatment offered to manic-depressive patients used to be psychoanalysis and maintenance ECT (Geoghegan 1949). The former was unfortunately not helpful and the latter effective but not favored by patients. Lithium was initially used only for the management of acute mania.

Since 1956 however, anecdotal observations started emerging about other possible benefits of lithium. Schou (1956) reported an observation of a manic patient who subsequently stopped having both manic and depressive recurrences when he maintained lithium during the free intervals. Beneficial action against depressions was also mentioned by Vojtechovsky (1957). Hartigan (1963) and Baastrup (1964) similarly noted that patients maintained on lithium had a marked reduction of both types of recurrence.

Baastrup and Schou (1967) then carried out a longitudinal study of patients with many previous episodes of illness. Patients with both bipolar and unipolar disorder were involved. The analyses indicated that recurrences occurred in patients significantly less frequently during lithium treatment than before such treatment, or even disappeared completely. Schou, Angst and I then decided to collaborate and to use a "mirror-image" design, utilizing the ample information we had about the previous course of illness of these manic-depressive patients. Against marked editorial resistance, our joined prospective observations on 250 lithium-treated patients were eventually published in the *British Journal of Psychiatry*. Together with clinical reports published earlier, an ample body of similar observations was emerging and demonstrating lithium as a useful drug in the treatment of manic-depressive illness.

Opposition against such interpretation emerged quickly, however, and among experts, views about the issue became sharply divided. Some psychiatrists expressed strong support for lithium prophylaxis, based on their own clinical experience. Others disagreed. On methodological grounds, Blackwell and Shepherd (1968) concluded that the claims for prophylactic efficacy were just a myth, supported by faulty evidence. They raised several critical points; their main objections were, first, a bias due to the open, non-blind evaluation of the recurrences and, second, a statistical approach which in their opinion weighted the facts in favor of the hypothesis.

But the objections could be effectively counteracted only by a tightly designed double-blind evaluation.

Barry Blackwell's invaluable contribution

Before commenting on these methodological disagreements, I want to express my gratefulness to Barry Blackwell. Even though his objections were incorrect, it was an invaluable service to moving ahead. Had it not been for his widely quoted procedural condemnation, I really wonder if lithium would now be in clinical practice all over the world. The national regulatory

bodies insist on double-blind tests. It was Blackwell's somewhat sarcastic, sharp, articulate arguments that made a strong impression and eventually forced the randomized double-blind trial. My feelings of gratefulness may not have been quite the same then but in hindsight they are strong, and I have expressed them repeatedly publicly.

When Mogens Schou, Jules Angst and I completed the replication study published in the *British Journal of Psychiatry* (1970), the last thing on our minds was to switch any of these patients to placebo. Many of them had suffered from severe, frequently recurrent mood disorder, had been hospitalized numerous times and badly incapacitated by their illness. On lithium they were stable for the first time and it seemed not only unethical but also unimaginable to ask them to stop it in order to participate in a clinical trial with placebo.

Before they went on lithium, I followed my patients for up to six years, failing to stop their depressions and manias. I could not imagine putting them and their families through the same misery again. In addition, in most of these patients the effect of lithium stabilization was so convincing, so different from the previous course of illness, that to use placebo just to prove that they again relapse appeared unethical and redundant. Furthermore, the Swiss and Czech findings were already an independent replication of the earlier Danish findings.

Finally, our analysis also indicated that the criticism aimed at our methodology was not correct. Barry Blackwell and Michael Shepherd raised two main methodological objections against the findings: that the marked recurrence reduction the patients experienced was to be expected naturally – that it was the result of a "regression to the mean" – and that the observations were not made blindly and thus biased by enthusiasm.

As for the duo's first protestation, the patients experienced frequent episodes qualifying them to enter the trial. Blackwell and Shepherd felt that the less frequent episodes that followed were the result of the recurrence frequency regressing back to a mean of lower value. To demonstrate their point, they quoted Saran's (1968) data of 13 patients who entered the follow-up with frequent episodes but lost that frequency with the passage of time.

But the problem was that, because of the small number of patients, Saran's example was neither representative nor applicable to the problem. Ottoson and Issakson (1969) and Laurel and Ottosson (1968) showed that the "mirror image" design is justified. In a sufficiently large sample of patients not receiving maintenance treatment the mean frequency of recurrences in the past becomes replicated in the future.

In essence, the individual clinical course of manic-depressive illness is capricious, overall seemingly random. Given this capriciousness, in a small group of patients the future frequency of recurrences will vary in any direction. It may decrease, as it did for Saran's 13 patients, it may increase, or it may remain about the same. But to obtain a predictable, anticipated mean frequency, one requires a sufficiently sizable cohort, as was the case in our open trials with a "mirror" design (1994).

As for the Blackwell and Shepherd's second objection – biased open assessment – blind evaluation is often very important but not a panacea. As Schou later demonstrated (1992), the results from long-term clinical trials of lithium were well comparable, regardless of whether the evaluations were carried out blindly or openly. Obviously, if one were evaluating the effects of an anxiolytic in neurotic patients, the placebo effect and bias would usually play a huge role. Double-blind arrangement would be indispensable. Blackwell illustrated clearly the relationship between observer enthusiasm and treatment outcome earlier with MAOI inhibitors (Blackwell and Taylor 1967).

But in a maintenance treatment of manic-depressive patients the task is markedly different: to assess if a patient who was previously symptom-free, develops in a free interval an acute manic or depressive episode. If in this task there would be large, systematic discrepancies between different psychiatrists with a similar training, we could forget psychiatry altogether.

Parenthetically, the biases of the involved investigators were actually markedly different, and not all positive. Schou and Baastrup were openly enthusiastic, because of their previous promising observations. Jules Angst appeared curious but neutral as to the expected outcome. And my previous attempts to prevent the recurrences of manic-depressive illness were so dismal (Grof and Vinao 1996, 1969), that I did not believe anything could work preventatively. As I wrote earlier, I was hoping to prove Schou wrong. Yet, despite our different preconceptions, our results with long-term lithium treatment were comparable.

Lithium's efficacy was subsequently proven in a trial that employed the design with blind evaluation and randomization (Baastrup, Poulsen and Schou 1970). As to the ethical concerns, using sequential analysis minimized the number of patients receiving placebo. Using sequential analysis can markedly reduce the number of patients needed to reach a statistically significant difference by utilizing, in addition, the probability hidden in the sequence in which the observations come in.

In this manner it became possible to complete the double-blind trial within six months and with a minimum of patients having been given placebo. All of the patients who became ill again were those switched to placebo, none of the lithium patients experienced recurrences during the same time.

Methodology in diapers

I fully concur with Barry Blackwell that one of the main reasons for our disagreements was the fact that the methodology of maintenance trials in bipolar disorders was in diapers then. In fact, we were developing the methodology while proceeding with the studies (Grof 1970).

As I read Barry Blackwell's "autopsy" I felt there were good reasons why we could not and cannot see lithium treatment in quite the same light. Our background, professional careers, experience and interests were different. As I understand it, his central interests were clinical trials, psychopharmacology and pharmacology. He was frustrated by many methodologically inadequate studies in the past and did not want to see another shabby study confusing psychiatrists. And, after his critique of lithium, he moved on to the industry and then academic and clinical practice. He seemed more interested in anxiety states than in following manic-depressive patients (Blackwell 2014).

We, on the other hand, prior to lithium studied the natural course of manic-depressive illness in hundreds of patients (Angst 1969; Angst, Dittrich and Grof 1969). Since the heated lithium debate, I have treated more than thousand patients with lithium, some of them up to 40 years and researched who does respond. With experience being so different, even now Barry Blackwell and my evaluation of lithium cannot be the same.

The efficacy of lithium is neither a myth nor imipramine-like

Barry Blackwell feels that, after the initial trials, uncertainty about Lithium's efficacy lingered until later studies published in 1984. He singles out Prien, Kupfer, Mansky et al. (1984), a double-blind trial carried out mainly in the US VA hospitals. The results are interpreted as indicating that in bipolar patients imipramine is better in cases of mild depressions and lithium in more severe cases. To claim that the efficacy of lithium is comparable to imipramine requires disregarding fully the rest of the published evidence. Such assertion seems to me idiosyncratic, neglecting the existing regulatory decisions, numerous clinical trials and expert consensus.

There is a body of double-blind clinical investigations together demonstrating prophylactic efficacy of lithium both against manias and depressions (Schou 1994; Coppen, Peet, Bailey et al. 1973), trials that have dealt well with Blackwell's methodological objections. On this basis, by the early 1970s lithium was approved for long-term treatment in most Western countries by regulatory agencies requiring solid double-blind evidence. Expert committees that have produced more than 25 guidelines for the treatment of bipolar disorder now quote lithium trials as the best (class I) evidence for efficacy. Despite Prien's study, imipramine is nowhere recommended for long-term treatment of bipolar disorders.

Prien's findings only can be interpreted if they are placed in the context of what had happened between 1968 and 1984 with diagnosing mood disorders. Manic-depressive illness was transforming into much larger and more heterogeneous "bipolar spectrum disorders." As lithium treatment had a striking success in patients with *typical* manic-depressive illness, the diagnostic fashion for mood disorders broadened (Baldessarini 1970; Grof and Fox 1987) and in "bipolar spectrum" disorders many patients with mood-incongruent symptoms and multiple comorbidities were included. In addition to manic-depressive illness the experimenting with lithium also expanded to other indications: schizoaffective conditions, cycloid psychoses, aggressive states, alcoholism, potentiation of antidepressants and several other situations.

But lithium prophylaxis is the treatment of choice only for what used to be "manic-depressive illness": in essence, remitting, episodically recurring bipolar and unipolar disorders. It may also be of partial help in other conditions, but the effect is quantitatively and qualitatively different: for example, one will see low efficacy and intense rebound after discontinuation. The diagnosis of manic-depressive illness used to require, among others, the exclusion of mood incongruent psychotic symptoms, the exclusion of other psychiatric diagnoses (i.e., exclusion of comorbidity) and the presence of episodic course.

This development created a very interesting situation. Recent studies have shown that bipolar disorder is now often underdiagnosed, particularly in recurrent depression. At the same time, as bipolar diagnosis is now given simply on the basis of a symptom set, without further analysis and exclusions, it is also grossly overused instead of other diagnoses. As a result, the bipolar spectrum disorder has become very fashionable and highly prevalent, but the classical lithium responsive manic-depressive patients are only a minority subgroup.

For whatever it's worth, while the Prien, Kupfer, Mansky et al. study was going on, two experienced American colleagues who knew my interest in lithium and participated in the study contacted me. They were very critical of the patient selection, partly due to the population of VA hospitals, and warned me not to believe the findings once the study is completed. In hindsight, the Prien, Kupfer, Mansky et al. study can hardly be considered the main pillar for the evaluation of lithium's usefulness.

Wide Acceptance and Narrow Opposition

Lithium treatment for bipolar disorder has gradually been accepted in most countries of the world, including the Third World countries. Lithium is now available as an effective mood stabilizer worldwide, but its use is geographically uneven. It should help in the Third World that lithium is inexpensive, particularly in comparison with new putative stabilizers.

But repeated questioning lithium's efficacy does happen and comes particularly from those who had been using lithium outside of the established evidence and in naturalistic studies with looser diagnosing and monitoring. But careful analyses have shown that lithium remains effective for patients with a clinical profile for which it was proven effective in the first place (Berghöfer, Alda, Adli et al. 2013).

Despite overwhelming evidence of the efficacy in typical manic-depressive cases, continuing debates about lithium are likely to occur between opponents who incorrectly believe that they are discussing the same issue but have used lithium in other bipolar types. Unfortunately, the correct evaluation of the outcome of stabilizing treatment in recurrent mood disorders is much more challenging than one would assume. Capricious course, fluctuating compliance with medication and a varying speed of stabilization all make it difficult to evaluate the relationship between the medication and a changed course of illness in any individual patient. Bipolar disorders have distinct subtypes responding preferentially to different mood stabilizers and lithium offers a variety of markedly different benefits to patients outside the classical manic-depressive illness (Grof 1998, 2003).

Squall

I do not know if Barry Blackwell really believes – as he seems to indicate in his writing – that the effect of lithium treatment on bipolar disorders is indeed comparable to the effect of

imipramine, or whether this is just another expression of his mastery of hyperbole. But we certainly do approach this issue from different angles.

When I think of lithium stabilization, my thinking is unavoidably colored by my experience of treating many bipolar patients for more than five decades. Before lithium treatment we lost every year several patients to suicide and the life of those who continued living was marred by their illness: the impact of frequent episodes of manias, depressions and hospitalizations and the influence on their families and professional life. Since we have been using lithium, the situation changed dramatically and these problems have been minimized, and often completely eradicated.

For Mogens Schou the initial heated debates were stressful. He was a very compassionate physician and switching stabilized patients from lithium to placebo troubled him greatly. He was also an extremely meticulous researcher. The possibility raised by Blackwell and Shepherd that he may have overlooked something important in methodology bothered him very much. He was extremely careful, as he kept moving between his laboratory and his clinical investigations.

The accusation of biased observation was not easy for him to swallow. There was some irony in blaming him for not having carried out the observations double-blind as he, in fact, performed the first double-blind study in psychopharmacology 15 years earlier. Particularly unfortunate was, I thought, Michael Shepherd's criticism ad hominem: he repeatedly stressed that Mogens Schou was a biased enthusiast because Mogens' brother's depressions responded well to lithium and, replying to questions, he never publicly conceded that lithium works.

From what Barry Blackwell has written about his professional life (Blackwell 2014), his professional interests have been different than ours and he worked more along different lines. His critique of lithium was an important but a relatively short-lived involvement and reflected more his interest in methodology and history of clinical trials than in the treatment of bipolar patients. Thus, I may be biased in favor of lithium but from his text I tend to conclude that he underestimates the helpfulness of lithium treatment and oversimplifies its use in bipolar disorders. Nevertheless, as I mentioned, in hindsight I see enormous value of his critique during the early days of lithium's clinical trials.

Impact of Lithium Treatment on Psychiatry

Up until 1967 no medication had seemed capable of averting recurrences of affective disorders; therefore, only acute episodes had been treated. The introduction of long-term lithium

treatment, lithium prophylaxis, changed things radically. From a practical point of view, it was primarily lithium's ability to prevent recurrences that made an impression. For research the introduction of lithium was a major stimulus for neurobiology, demonstrating that a simple element can produce major neurobiological changes. Lithium became the focus of attention of pharmacologists, biochemists, physiologists, psychiatrists, psychologists and many others. It was probably the advent of lithium therapy that made psychiatric research truly interdisciplinary. Research on all aspects of the affective disorders has been greatly stimulated by the demonstration of the effectiveness of lithium in the treatment of these conditions.

For academic psychiatry the acceptance of lithium treatment led to the important recognition that mood disorders are much more common than previously presumed and that the existing classification systems must be reconsidered. As the history of the past four decades has shown, lithium therapy has made a significant contribution to modern psychiatry, both in relation to its specific uses in alleviating recurrent endogenous affective disorders and in stimulating psychiatric research and conceptual thinking.

References:

Angst J, Weis P, Grof P, Baastrup PC, Schou M. Lithium prophylaxis in recurrent affective disorders. Br J Psychiatry. 1970;116:604-14.

Angst J BP, Grof P, Hippius H, Poldinger W, Weiss P. Clinical course of affective disorders. Psychiatry. 1969;76:489-500.

Angst J, Dittrich A, Grof P. Course of endogenous affective psychoses and its modification by prophylactic administration of imipramine and lithium. Int Pharmacopsychiat. 1969;2:1-11.

Baastrup PC. The use of lithium in manic-depressive psychosis. Compr Psychiatry. 1964;5:396-408.

Baastrup PC, Schou M. Lithium as a prophylactic agent: its effect against recurrent depressions and manic-depressive psychosis. Arch Gen Psychiatry. 1967;16(2):162-72.

Baastrup PC, Poulsen JC, Schou M. Prophylactic lithium: Double blind discontinuation in manic-depressive and recurrent-depressive disorders. The Lancet. 1970;2:326.

Baldessarini RJ. Frequency of diagnoses of schizophrenia versus affective disorders from 1944 to 1968. American Journal of Psychiatry. 1970;127:59-63.

Berghöfer A, Alda M, Adli M, Baethge C, Bauer M, Bschor T, Grof P, Müller-Oerlinghausen B, Rybakowski JK, Suwalska A, Pfennig A. Stability of lithium treatment in bipolar disorder - long-term follow-up of 346 patients. Int J Bipolar Disord. 2013;1:11.

Blackwell B. The Lithium Controversy: An Historical Autopsy. inhn.org.controversies. June 19, 2014.

Blackwell B, Shepherd M. Prophylactic lithium: Another therapeutic myth? An examination of the evidence to date. The Lancet. 1968;I:968-971.

Blackwell B, Taylor D. An Operational Evaluation of MAOI. Proceedings Royal soc. Med. 1967.

Coppen A, Peet M, Bailey J, Noguera R, Burns BH, Swani MS, Maggs R, Gardner R. Double-blind and open studies of lithium prophylaxis in affective disorders. Psychiatr Neurol Neurosurg. 1973;76:501-10.

Geoghegan JS, Stevenson GH. Prophylactic electroshock. Am J Psychiatry. 1949;105:494-6.

Grof P, Müller-Oerlinghausen B. A critical appraisal of lithium's efficacy and effectiveness: the last 60 years. Bipolar Disorders. 2009;11:10-19.

Grof P. Designing long-term clinical trials in affective disorders. J Affect Disord. 1994;30:243-55.

Grof P, Vinao O. Maintenance and prophylactic imipramine doses in recurrent depressions. Activitas Nervosa Superior. 1966;8:384-5.

Grof P, Vinao O. Comparison of various prophylactic procedures in affective psychoses. In: Das Depressive Syndrom, Hippius H, Selbach H, editors. Berlin 1969. pp. 101-4.

Grof P, Schou M, Angst J, Baastrup PC, Weis P. Methodological problems of prophylactic trials in recurrent affective disorders. Br J Psychiatry. 1970;116:599-603.

Grof P, Fox D. Admission rates and lithium therapy (letter to the editor). British Journal of Psychiatry. 1987;150:264-265.

Grof P. Has the effectiveness of lithium changed? Impact of the variety of lithium's effects. Neuropsychopharmacology. 1998;19:183-8.

Grof P. Selecting effective long-term treatment for bipolar patients: monotherapy and combinations. J Clin Psychiatry. 2003;64 Suppl. 5:53-61.

Hajek T, Cullis J, Novak T, Kopecek M, Hoschl C, Blagdon R, O'Donovan C, Bauer M, Young LT, Macqueen G, Alda M. Hippocampal volumes in bipolar disorders: opposing effects of illness burden and lithium treatment. Bipolar Disord. 2012;14:261-70.

Hartigan GP. The use of lithium salts in affective disorders. Brit J Psychiat. 1963;109:810-814.

Isaksson A, Ottosson JO, Perris C: Methodologische Aspekte der forschung uber prophylaktische Behandlung bei affektiven Psychosen. In: Das Depressive Syndrom Hippius H, Selbach H, editors. Berlin. 1969, pp. 561-574.

Laurell B, Ottosson JO. Prophylactic Lithium? The Lancet. 1968;2:1245.

Müller-Oerlinghausen B, Ahrens B, Volk J, Grof P, Grof E, Schou M, Vestergaard P, Lenz G, Simhandl C, Thau K, et al. Reduced mortality of manic-depressive patients in long-term lithium treatment: an international collaborative study by IGSLI. Psychiatry Res. 1991;36:329-31.

Prien R, Kupfer DJ, Mansky PA, Small JG, Tuason VB, Voss CB, Johnson WE. Drug therapy in the prevention of recurrences in unipolar and bipolar affective disorders: report of the NIMH collaborative study group comparing lithium carbonate, imipramine, and a lithium carbonate-imipramine combination. Archives of General Psychiatry. 1984;41:1096-1104.

Quiroz J.A.C.; Drevets, W.C.; Henter, I.D.; Manji, H.K.: Mood Disorders (Chapter 3). In: Translational Neuroscience, Barrett JEC, Williams M, editors. Cambridge University Press; 2012. pp. 27-69.

Schou M. Litiumterapi ved mani: Praktiske retningslinier. Nord Med. 1956;55:790-4.

Saran BM. Prophylactic lithium? Lancet 1968 Aug 3; 2(7562):284-5.

Schou M: Phases in the development of lithium treatment in psychiatry. In: The Neurosciences: Paths of Discovery II, Samson FAG, editor. Birkhäuser: Boston, Basel, Berlin. 1992. pp. 148-66.

Vojtechovsky M: Zkusenosti s lecbou solemi lithia. Problemy Psychiatrie v Praxi a ve Vyzkumu. Prague, Czechoslovak Medical Press 1957. pp. 216-24.

Jules Angst's comments

Studies on the long-term natural history of mood disorders

Knowledge of the course of mood disorders is essential when deciding whether long-term prophylactic medication is justified. This is especially the case if the natural history of a disorder shows not spontaneous improvement but rather persistent recurrence or even an increase of episodes, reflected by shortening cycle lengths. A cycle is defined as an episode plus the subsequent interval, i.e., the time between the onset of two subsequent episodes.

In 1967 Angst and Weis, investigating a group of 375 subsequent hospital admissions of patients with mood disorders in Zurich, found a log-normal distribution of episode and cycle lengths in four subgroups (125 recurrent depression, 117 involutional melancholia, 45 bipolar and

85 schizo-affective psychoses). This signified that studies should no longer base on arithmetic means. More important, however, was that the longitudinal analysis of cycle lengths showed a clear acceleration of recurrences with an increasing number of episodes. This was most marked in patients with bipolar disorder, followed by schizo-affective disorder and was lowest in patients with depressive disorders.

These findings were reproduced in 386 patients from a further three centers (Basle, Berlin and Landeck) by Angst, Grof, Hippius and Weis in 1968. In 701 patients with bipolar disorder and 988 patients with recurrent major depression a progressive shortening of cycles correlated with age at onset, age and number of episodes over 20 years (the subsequent cycle was about 10% shorter than the previous one). Despite the clear finding of an increasing recurrence risk by shortening of cycles, in our statistical testing of the long-term effect of prophylactic treatment we applied the very conservative *mirror model*, which assumes, as the zero hypothesis, merely an equal occurrence of the number of episodes before and under treatment during identical, individual observation periods.

Studies on prophylactic treatment of mood disorders with imipramine and lithium

A first study by Angst, Dittrich and Grof in 1969 dealt with patients treated with imipramine (N=63) or lithium (N=91) in Prague and Zurich. Statistically the mirror model was applied with Wilcoxon signed rank tests. Under imipramine there was a significant deterioration in the course of depression during the second intra-individual period, whereas lithium showed a positive effect in bipolar disorders ($p<.025$) and a trend to an effect in recurrent depression or involutional melancholia ($p<.07$). The negative and positive effects of the two drugs were comparable and significant in both samples (Prague and Zurich). This was the first statistically based demonstration of the efficacy of lithium in treating mood disorders.

In a larger analysis of lithium data undertaken in Glostrup, (DK), Prague and Zurich (Angst, Weis, Grof, Baastrup and Schou 1970) equal observation periods (before and under treatment) were again compared with regard to hospital admissions and number of episodes before and during lithium prophylaxis. In all three centers there was a reproducible significant decrease in the number of episodes ($p<.001$) during the lithium period. Taking all 244 patients together, there were significantly fewer hospital admissions for patients with manic-depressive disorders ($p<.001$), recurrent depression ($p<.002$) and schizo-affective disorders ($p<.01$). The average

observation periods of the three groups compared were 2x38.5 months for bipolar disorder, 2x26.7 months for recurrent depression and 2x28.1 months for schizo-affective psychoses.

Thus, in contrast to the repeatedly confirmed deterioration of the spontaneous course of mood disorders, a significant improvement was achieved with lithium but not so with imipramine.

Personal reminiscences

Ervin Varga from Budapest and I worked under Michael Shepherd for a few months, researching 981 records of patients treated for depression in the Maudsley Hospital. At that time my monograph showing the differences between unipolar depression, bipolar disorder and schizo-affective disorder (published in 1966) had already been accepted for print. Bleuler, Strömgren and Aubrey Lewis agreed that the results could not be true, but fortunately Eliot Slater believed in the correctness of the findings. Strömgren and Bleuler changed their minds after Perris published similar results. Michael Shepherd had by self-admission never treated a patient with lithium and I would be interested to know whether Blackwell himself had done so at the time of their joint article in the Lancet.

During these years we held annual meetings of the International group of studies of affective disorders (IGSAD) founded by Ottosson, Perris, Winokur and Angst. I invited Michael Shepherd to join the group. Some members (Christian Baastrup, Mogens Schou, Max Hamilton, Martin Roth, Paul Grof, Jules Angst, etc.) discussed the ethical and feasibility problems of a placebo-controlled study on lithium and decided finally on the design of a cessation study carried out by Glostrup (Baastrup, Poulsen, Schou et al., 1970).

The positive results of this study were first presented to an IGSAD meeting and were the subject of intense debate. Michael Shepherd remained silent during the discussion and when asked for his opinion, replied "no comment." At the end of that year, I asked him to retire from the group, which he did; we remained on very friendly terms for the rest of his life.

References:

Angst J, Weis P. Periodicity of depressive psychoses. In: Brill H, Cole JO, Deniker P, Hippius H, Bradley PB, editors. Neuropsychopharmacology. Proceedings of the Fifth International Congress of the Collegium Internationale Neuropsychopharmacologicum. Washington D.C. 1966, 28-31

March 1966. International Congress Series No.129. Excerpta Medica Foundation 1967, pp 703-10.

Angst J, Grof P, Hippius H, Poeldinger W, Weis P. La psychose maniaco-dépressive est-elle périodique ou intermittente? In: Cycles biologiques et psychiatrie, De Ajuriaguerra J, editor. Paris, Masson, 1968, pp 339-51.

Angst J, Dittrich A, Grof P. Course of endogenous affective psychoses and its modification by prophylactic administration of Imipramine and Lithium. Int Pharmacopsychiat 1969;2:1-11.

Angst J, Weis P, Grof P, Baastrup PC, Schou M. Lithium prophylaxis in recurrent affective disorders. Br J Psychiatry 1970;116:604-14.

Baastrup PC, Poulsen JC, Schou M, Thomsen K, Amdisen A. Prophylactic lithium: Double-blind discontinuation in manic- depressive and recurrent-depressive disorders. Lancet 1970;2:326-30.

January 22, 2015

Barry Blackwell's reply to Paul Grof's and Jules Angst's comment

I thank Paul Grof for the kindness and generosity of his comments and must confirm what both he and Jules Angst suggest was my youthful inexperience at the time of the controversial "Prophylactic Lithium" article in the Lancet co-authored with Michael Shepherd. Much of my residency training (1962-1967) was preoccupied with human and pharmacology research on the interaction of MAO inhibitors and tyramine containing foods so I had virtually no practical experience with lithium.

It is also accurate that different career patterns have colored our opinions of the research and its practical implications. Both Drs. Grof and Angst have devoted significant portions of their careers to sophisticated research on the natural history and drug treatment of the bipolar affective disorders in large populations of patients. My own career trajectory has been entirely different, devoted to a wide spectrum of interests in pharmacology, psychosomatic medicine, medical education and specific topics such as patient compliance, homelessness, chronic pain, physician career development and administration of two academic departments. As a result, my continuing interest and knowledge in the arena of bipolar disorder became that of a journeyman (albeit

academic) psychiatrist with only a modest involvement in everyday clinical practice and a patchy knowledge of the evolving literature.

From my personal experience and those of colleagues I did learn how the depressive component of bipolar disorders often persists in subdued form despite lithium and is difficult to treat with imipramine (or anything else) without the risk of aggravating manic symptoms. So, this debate in enlivened by the question of how much a body of academic research knowledge can be reliably and usefully transferred as relevant to everyday clinical practice.

How much of value I may have missed in this search is problematic. Paul Grof's bibliography confirms his comment on the lengthy lapse in general interest and research on lithium's effects in bipolar disorder. His list of references covers over seven decades (1940-2015) and of his 64 citations exactly half (32) appeared during the single decade (1961-1970) when this debate erupted. No other decade has more than three citations until 2010, since when five new publications appear of which Paul is a co-author on three.

Review of this and Jules Angst's literature suggests that although our differences are largely reconciled lingering issues are worthy of debate.

It is a fact that one methodological concern raised by Shepherd and me related to potential bias due to lack of a double-blind and also true that we underestimated the understandable concerns about safety and suicide that later resulted in imaginative alternative research designs. A second was the possibility of statistical regression to the mean. A third, and perhaps major concern, was the heterogeneity of the patient sample. Both of these latter two concerns were elegantly displayed by the article's graphic portrayal of episodes of illness and remission including recurrent manic, depressive and mixed forms. At a time when imipramine had established its efficacy as an antidepressant in single episodes of unipolar depression we questioned if it also might have a prophylactic effect. We tested this hypothesis using Baastrup and Schou's statistical model in a sample of recurrent unipolar depressed patients treated with imipramine from the Maudsley Hospital data base and found it to be confirmed. This colored our conclusion that lithium was unlikely to be prophylactic for the entire spectrum of bipolar disorders. In retrospect our overboard response to this finding might belong in the category known as "throwing out the baby with the bathwater."

Much of Paul Grof's and Jules Angst's ongoing research has been devoted to a more specific clarification of what types of bipolar spectrum disorder respond in which manner to

lithium, imipramine and other mood stabilizers. Grof notes (p.24) that, "recent studies have shown that bipolar disorder is now often undiagnosed, particularly in recurrent depression." This conclusion may have been embedded in Baastrup and Schou's original study and unkindly labeled by us as "bias." My current assumption based on Paul Grof's information is that Schou's brother failed both ECT and imipramine before responding dramatically and persistently to lithium despite being previously considered to suffer from recurrent unipolar depression. Perhaps this conviction was reflected in their diverse patient population and claim for ubiquitous benefit across the spectrum of disorders. Schou's late life interest in this topic suggests he might have been seeking for subtle manifestations of hypomania between episodes of severe depression that would indicate a lithium responsive diathesis. Experienced clinician that Schou was, perhaps he was correct in this also.

Dr. Grof is dismissive of the Prien, Kupfer, Mansky et al. (1984) study and the weight it accords in support of imipramine's potential benefit in recurrent depressive disorder. He cites an impressive body of contradictory evidence which includes personal phone communications from two experienced American colleagues working with Prien who were, "very critical of the patient selection, partly due to the population of V.A. hospitals, and warned me not to believe the findings once the study was completed." The precise scientific basis for this conclusion is not revealed but the allegation is sadly reminiscent of the *ad hominem* feelings evoked by us in Schou's sensitive response to our better articulated concerns. I regret this.

Overall, I strongly agree with Grof and Angst that this kind of historical dissection can be beneficial to contemporary understanding of the evolution of neuropsychopharmacology. We all make mistakes and owning them may benefit posterity. Perhaps the lithium controversy belongs in the larger context pervading our entire field. The first two decades of clinical psychopharmacology were filled with expectations that we would "discover the right drug for the right patient." The story of lithium, the first of our truly psychotropic drugs, so well portrayed by Paul and Jules, shows how far we have come but have yet to go in achieving that end. As Dr. Grof notes, expert committees around the world have produced 25 guidelines for the treatment of bipolar disorder despite which he notes the lax diagnostic practices, overuse and unrealistic expectations for lithium.

Despite the best efforts of dedicated researchers to define therapeutic specificity the general practitioners in our field (of which I was one) continue to operate on a "trial and error basis" when

selecting drug treatment for an individual patient. This is contributed to by our still incomplete knowledge of the natural history, genetic origins and phenotypic presentations of the disorders, an unhelpful DSM system of diagnosis, complicated by side effect sensitivity, drug interactions, differing drug profiles, variable compliance and misleading commercial mythologies regarding drug specificity.

If this sounds like "masterful hyperbole" (of which Paul Grof accuses me), please read our essay, "Sir Aubrey Lewis" (Goldberg, Blackwell and Taylor 2015) to better understand the differences between hyperbole (OED: deliberate exaggeration, not to be taken literally) and empiricism (OED: knowledge based on observation and experiment). The latter philosophy of science was the model in which I was trained, a style for which the Maudsley was both renowned and denigrated and of which our article, "Prophylactic Lithium: a Therapeutic Myth" remains, for all its faults, a paradigm.

This deconstruction of the lithium controversy brings to mind a final concern. At their inception more than half a century ago both the ACNP and the CINP established policies and memberships dedicated to translational research and dialog. This had dual implications: that basic science might illuminate clinical research while clinical research of the caliber conducted by Grof and Angst would translate to improved everyday diagnosis and treatment by practitioners in the general fields of psychiatry and medicine.

During my own residency training "Descriptive Psychiatry" was the prevailing idiom in European psychiatry – a dedicated interest in the nosology and natural history of mental disorders, illuminated by biological, psychological and social influences and insights although treatment options were sparse. As the new drugs appeared this interest survived initially but began to wane as the connection between clinical features and outcomes was influenced by discoveries, speculations and false hopes involving neurotransmitters, receptors, neural pathways, hormonal and genetic influences. The membership and interests of the ACNP began to tilt unevenly in the direction of neuroscientists (often with dual doctorates), the number of talented clinicians dwindled by attrition while clinical research and data analyses were increasingly usurped by industry. At the same time the NIMH withdrew from new drug research. The well-intentioned DSM nosology is capable, if scrupulously used, of sustaining interest in descriptive psychiatry and sophisticated biopsychosocial formulations but its multi-axial potential has been degraded to become primarily

an Axis 1 diagnosis for insurance purposes and ubiquitous use of the Not Otherwise Specified (NOS) categories.

I hope that this renaissance of interest in natural history, nosology and treatment outcomes in bipolar spectrum disorders, sparked by Grof and Angst's research, will have a wider influence on the future direction of our field.

Reference:

Goldberg D, Blackwell B, Taylor D. Sir Aubrey Lewis. Professor Sir Aubrey Lewis, the Maudsley Hospital & the Institute of Psychiatry. inhn.org.biographies. February 19, 2015.

February 5, 2015

June 11, 2020