Reshaping the carcinogenic risk assessment of medicines: international harmonisation for drug safety, industry/regulator efficiency or both?

John Abraham*, Tim Reed

Centre for Research in Health and Medicine (CRHaM), University of Sussex, UK

Abstract

The most significant institutional entity involved in the harmonisation of drug testing standards worldwide is the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), which comprises the three pharmaceutical industry associations and regulatory agencies of the EU, US and Japan. It is often claimed that such harmonisation will both accelerate the development and approval of new drugs and preserve safety standards, if not strengthen safety regimes. Drawing on extensive documentary research and interviews, this paper systematically examines whether the efforts by the ICH to improve industrial and regulatory efficiency by harmonising drug testing requirements is likely to raise, maintain or compromise safety standards in carcinogenic risk assessment of pharmaceuticals. The evidence suggests that, in the field of carcinogenicity testing, the ICH management of international harmonisation of medicines regulation is not achieving simultaneous improvements in safety standards and acceleration of drug development. Rather, the latter is being achieved at the expense of the former. Indeed, the ICH may be converting permissive regulatory practices of the past into new scientific standards for the future. These findings are significant as many expert scientific advisers to drug regulatory agencies seem to have accepted uncritically the conclusions reached by the ICH, which may affect a potential patient population of half a billion and tens of thousands of clinical trials.

*Corresponding author. School of Social Sciences, Arts E Building, University of Sussex, Falmer, Brighton BN1 9SN, UK. Tel.: +44-273-678883; fax: +44-273-62357.
E-mail address: j.w.abraham@sussex.ac.uk (J. Abraham).

0277-9536/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
PII: S 0277-9536(02)00339-8
has been a growth in regulatory workload and an increasing conflict with the protection of public health. Thus, as Morris (1997, p. 60) notes, international harmonisation of regulatory standards has become attractive for many regulatory agencies because, by agreeing to accept each others decisions under harmonisation, duplication of effort may be reduced and workload shared. This has been accentuated by the demands of patients groups for faster access to some drugs promising significant therapeutic benefits for high-profile life-threatening illnesses, such as AIDS (Epstein, 1996).

The most significant institutional entity involved in such harmonisation is the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), which comprises the three pharmaceutical industry associations and regulatory agencies of the EU, US and Japan. While the International Federation of Pharmaceutical Manufacturers’ Associations (IFPMA) acts as the secretariat of ICH, many of the key participants are industry and government scientists, who formed ‘expert working groups’ (EWGs) in areas of drug safety, quality and efficacy with a view to harmonising related regulatory standards across the three regions. These experts developed 37 ICH guidelines at four major ICH meetings in 1991, 1993, 1995 and 1997 (IFPMA, 2000).

The guidelines are of major significance because the industry and regulatory agencies in the EU, US and Japan have committed themselves to follow them. The final stage of the ICH procedure is the adoption of the guidelines into the domestic regulations of the EU, Japan and the US, after which the regulatory authorities are bound to implement them. Generally, regulatory authorities do not have the option of requiring higher safety standards than those agreed at ICH. Furthermore, these guidelines are also likely to influence the World Health Organisation (WHO) and regulatory agencies of Canada, Norway, Iceland and Liechtenstein, who were ‘observers’ at ICH meetings. Davis (1997, p. 145) argues that ‘if managed correctly’ international harmonisation of regulatory standards could reconcile the apparently contradictory pressures towards ‘readier access to drugs and more stringent safety requirements’ by both strengthening safety regimes and accelerating the approval of new drugs. In theory, therefore, international harmonisation does not necessarily produce a diminution or an improvement in drug safety standards because the outcome is crucially dependent on how the process is managed. While analytically sound, this view attaches no immediate significance for drug safety to the coincidence in timing of harmonisation and neo-liberal influences on the regulatory state. By contrast, sceptics of harmonisation emphasise its links with a deregulatory political climate around issues of safety. For example, Hodgkin (1996, p. 3) claims that the ICH process is ‘strongly biased towards getting new drugs on to the global market as quickly as possible’ rather than towards public health and safety. While, these perspectives have usefully highlighted the parameters of debate about international harmonisation and drug safety, there has been no systematic social scientific analysis of how management of the ICH process is actually influencing drug safety standards.

In this paper, therefore, we test (in a limited way) two related hypotheses: that ICH is being managed both to accelerate drug development/regulation and to maintain/raise safety standards; or that the management of ICH is accelerating drug development and regulation at the expense of drug safety standards. It follows logically that the validity of these hypotheses rests crucially on whether safety standards are being maintained or eroded. Specifically, our test is limited to an examination of how the ICH process is reshaping the regulatory standards for assessing the risk of inducing cancer (carcinogenicity) that new drugs may pose to patients. This is an important area because cancer remains one of the major killers of people in the industrialised world. A consequence of this focus is that our investigation touches on only one of ICH’s guidelines. Apart from its public health significance, the carcinogenicity guideline has been chosen because it is recognised by industry as having the greatest impact.

A major advantage of our narrow approach is that it permits an in-depth analysis which can take social science debate beyond the broad commentaries which have attended discussions of international harmonisation of medicines to date. In particular, the socio-technical relationships between harmonisation and the regulatory definitions of drug safety and carcinogenic risk to public health can be scrutinised effectively. Nevertheless, we also discuss why our findings may be part of a more general trend within ICH regarding drug safety standards. It should be noted, however, that in developing its guideline on carcinogenicity testing, the ICH recognised ‘special situations where it may not be necessary to require carcinogenicity tests to be completed before marketing’, such as in the development of drugs to treat ‘life-threatening diseases, especially where no satisfactory alternative therapy exists’ (D’Arcy & Harron, 1996, p. 258). Thus, the ICH guideline on

1 Interview with Deputy Director of the Swedish drug regulatory authority, the Medical Products Agency, 5 February 1998.
2 Interview with Vice-President of Regulatory Affairs, Hoechst Marion Roussel, 16 January 1998.
3 Interview with Vice-President of International Affairs, IFPMA, 7 January 1998.
carcinogenicity testing is, in effect, a guideline for the testing of drugs for non-life-threatening diseases/illnesses.

When testing new drugs for carcinogenicity regulators and industrial scientists have had to rely on the extrapolation of results from non-human tests because most carcinogenic risks accelerate over the lifespan—70–80 years for humans—far too long for clinical trials (Schou 1992, p. 210). Since the late 1960s, two types of carcinogenicity studies have been developed: short-term in vitro mutagenicity tests; and lifespan in vivo studies in rodents (Hayashi, 1994, p. 291). The former involves adding the test drug to mammalian/human cells or to micro-organisms in glass dishes in order to see if the chemical alters/damages DNA, causing mutations, associated with carcinogenic risk (King, 1996, pp. 93–94). Pharmaceuticals found to damage DNA in vitro are known as genotoxic carcinogens. However, many cancer-inducing chemicals do so without damaging DNA as a primary biological activity (Purchase 1992). These non-genotoxic carcinogens are not detected by short-term mutagenicity tests (Ashby & Tennant, 1991). Hence, in addition, animal models of human carcinogenesis have been employed to screen for non-genotoxins by feeding rodents the test drug over their lifespan, usually between 18 and 24 months. At the end of the study, the incidence and nature of the tumours found among the rodents given the test drug are compared with those in a ‘control group’ which do not receive the drug. However, because of the problem of extrapolating findings to another species, and to humans in particular, the WHO (1969) recommended that lifespan carcinogenicity testing should be conducted in at least two species (typically mice and rats) prior to marketing approval. By the late 1970s, the regulatory agencies in North America, Western Europe and Japan accepted this view (Abraham, 1998).

Thus, in testing the above hypothesis via a case study of carcinogenic risk assessment, we also explore how changes to regulatory science are produced and their sociological relationships to political culture, neo-liberal ideology, industrial interests and public participation in policy. This provides an understanding of why the management of ICH took the shape that it did, and informs future work concerned with improving on the ICH’s construction of drug safety standards. We conclude our paper by putting forward an alternative way of locating carcinogenicity testing in the relationship between the regulatory science of risk assessment and public policy.

Methodology

The defining feature of the research method was a combination of documentary research complemented by targeted ‘informant’ interviews. The documentary research involved systematic searches of the relevant scientific, medical, regulatory and industry literature together with a comprehensive review of all the published proceedings of the first four ICH meetings from 1991 to 1997. English-language on-line searches of BIDS (Web of Science) and Medline for the years 1990–1998 were conducted under all the relevant institutional descriptors and systematically matched with ICH, carcinogenicity testing, drug safety, drug testing and drug toxicology. These latter four were also matched alternately and systematically with ICH, pharmaceutical innovation and pharmaceutical R&D. In addition, the pharmaceutical trade publication, SCRIP, was searched manually from 1990 to 98 and all ICH guidelines and reports were collected.

Following the documentary research, interviews were conducted with individuals from organisations whose work was, or could be, relevant to the relationship between ICH, drug development, carcinogenic risk assessment and safety standards. For practical reasons, the interview research was confined to Europe. As far as possible, individuals were chosen for interview because of their knowledge about the ICH process. In this sense our sample of interviewees was chosen mainly as informants, rather than on any quantitative criterion of representativeness.

For example, within industry, we interviewed representatives of the Centre for Medicines Research (CMR), the IFPMA, and 11 major pharmaceutical companies involved with ICH to clarify our understanding of the ICH process from our documentary research. Key regulators involved in the ICH process, such as representatives of the European Medicines Evaluation Agency (EMEA) and its expert Committee for Proprietary Medicinal Products (CPMP), were also selected and interviewed with a similar rationale, as were the ‘observers’, the WHO and the European Free Trade Area (EFTA). However, we were also interested in what national regulators and other non-governmental organisations outside the ICH process thought about the ICH’s work on drug safety and carcinogenic risk assessment. To this end we sampled national regulators from the Swedish Medical Products Agency (MPA), the UK Committee on Safety of Medicines (CSM) and the UK Medicines Control Agency (MCA), medical representatives from the British Medical Association (BMA) and the Committee for the International Organisation of Medical Sciences (CIOMS), and the public health advocates from the UK Consumers’ Association (CA). In total, 32 informants were interviewed at 25 meetings during 1997 and 1998. Seventeen were from industry, nine were from regulatory agencies or their expert advisers and six were from the other organisations, giving a response rate of 60, 90 and 100 per cent, respectively. All interviews were tape-recorded and
anonymised in accordance with the wishes of the informants. The interviews were semi-structured and tailored specifically for particular subjects.

The critique of lifespan studies in rodents

In the overall non-clinical safety evaluation of a new drug, the lifespan carcinogenicity study in rodents consumes the most time and resources; taking over 3 years and costing about one million US dollars per compound tested. The results of these studies can have a major impact on the approval of a product (Contrera, 1996). According to Monro (1996, p. 262), a representative of the US Pharmaceutical Research and Manufacturers Association (PhRMA), such a lengthy operation can be a major deterrent to some drug development programmes.

In the 1980s the Association of the British Pharmaceutical Industry (ABPI) established the CMR to ‘monitor the regulatory process in different countries in order to provide data which will support recommendations to expedite this process’ (CMR, 1995, p. 3). The CMR and other elements within the pharmaceutical industry questioned the relevance of findings from, and the predictive value of, rodent lifespan carcinogenicity tests for assessing human risk of cancer as many chemicals found to provoke cancer in rodents did not seem to cause cancer in humans (Bentley, Baier & Krug, 1992). In particular, it was argued that carcinogenicity testing should focus more attention on the mechanisms of cancer induction, as these might be species-specific and the CMR formulated a short-term goal of reducing the regulatory standards to just one lifespan carcinogenicity study in one rodent species (Mulliger, 1997, p. 16).

It is certainly the case that for some in industry this was linked to commercial interests. As one manager noted: ‘obviously the industry has an interest in trying to reduce the amount of time one spends in developing products and the best way of doing that is to reduce the requirements for long-term animal testing .... pharmaceutical companies have to get their products on the market quickly and get some sort of return, otherwise they can’t keep their shareholders happy’.4 Furthermore, interviews revealed that many in industry have no hesitation in viewing ICH as an integral part of accelerating drug development and regulation.

Consequently, carcinogenicity testing of new drugs became a significant topic for ICH. Although the remit of the ICH was initially to harmonise inconsistent regulatory standards, in the area of carcinogenicity testing, where international regulatory standards were the same, it sought a ‘consensus reducing requirements for carcinogenicity studies at ICH’ (Anon, 1996, p. 259). The implications of such reductions for public health are complex, but the simple message from one industrial scientist that ‘the less we are doing in toxicology, the higher is the risk to humans at the end of the day’ signals that savings in drug testing may have adverse consequences for patients’ health.

While a reduction in the time and resources required for non-clinical drug safety testing was clearly one motivation for ICH’s review, the critique of lifespan carcinogenicity studies conducted in public focused on their scientific and regulatory validity. At ICH the critics of lifespan carcinogenicity studies contended that several of the mechanisms elucidated for non-genotoxic rodent carcinogenesis caused by pharmaceutical substances do not operate in humans under the conditions of drug exposure (Monro, 1992). They claimed that scientific understanding of the mechanisms involved in such carcinogenesis has increased greatly, so that it is possible to distinguish those mechanisms that appear to be specific to the animal model from those that could be reasonably expected to have trans-species, and hence human, relevance (Monro, 1994).

According to Monro (1994), the US National Toxicology Program (NTP) database did not show a good correspondence between the tissue sites of occurrence of human cancer and the target sites for chemical carcinogens in rodents. Cancers of the colon and rectum are rare in rodents, but common in humans, and vice versa for liver and lung cancers (excluding tobacco-related cases). Thus, Monro condemned rodent lifespan carcinogenicity tests for generating too many false positives, that is, positive carcinogenicity results in rodents for drugs that are not carcinogenic to humans.

These critics give the impression that scientific progress in understanding the mechanisms of cancer have undermined the validity and relevance of lifespan carcinogenicity testing in rodents for human hazard assessment. However, many scientists have expressed an opposing view that the relevance of toxicity findings for humans are increased if the same effect is observed in two or more mammalian species, rather than in one (Parkinson & Grasso, 1993; Zbinden, 1987, pp. 51–52). Similarly, Schou (1992, p. 210) has acknowledged that ‘it is generally agreed’ that the lifespan carcinogenicity study is the test which ‘gives the optimal answer to the question if a new drug presents a carcinogenic risk’. Furthermore, when Monro was asked, at ICH, how he could be sure that there are too many false positives in rodent lifespan carcinogenicity studies, he replied: ‘I am not certain whether there are too many [carcinogenic]

4 Interview with Director of Regulatory Affairs, Schering-Plough, 12 December 1997.
5 Interview with Director of Safety Assessment, Astra Charnwood, 18 December 1997.
6 Interview with Director of Toxicology, Astra, 6 February 1998.
activities ... some of these positives may as yet be unidentified human carcinogens’ (Ekman, 1996, p. 298).

All are agreed with Tennant (1993) that trans-species carcinogens should be distinguished from single-species carcinogens because the former pose a greater risk to human health. While critics of lifespan carcinogenicity testing in rodents argue that these distinctions can be made with short-term tests exploring the mechanisms of cancer induction, other scientists believe that the existence of trans-species carcinogens underlines the need for lifespan carcinogenicity testing in at least two animal species in order to identify them. Without lifespan carcinogenicity tests in two species there is a risk that some trans-species carcinogens will not be identified, and that regulators will not have a negative result confirmed in a second species. For these reasons, scientists at the US National Institute of Environmental Health Sciences (NIEHS) and the NTP recommended that lifespan carcinogenicity studies should be conducted in at least two animal species, and that ‘in the absence of adequate data on humans, it is biologically plausible and prudent to regard agents for which there is sufficient evidence of carcinogenicity in experimental animals as if they presented a carcinogenic risk to humans’ (Fung, Barrett, & Huff, 1995, p. 682).

It is clear from this review that, during the 1990s, there did not exist harmonious agreement about the value of rodent lifespan carcinogenicity testing of pharmaceuticals to drug safety regulation. Well into the 1990s, an FDA science symposium on carcinogenic evaluation of pharmaceuticals noted that ‘a vigorous debate is underway concerning the utility of 2-year rodent bio-assays to human health risk assessment, with strong opinions existing on all sides’ (Fenner-Crisp, 1996).

The attack on the mouse study

Despite the considerable scientific uncertainty about whether lifespan carcinogenicity studies should be conducted in one or two rodent species, as early as 1992, the ICH EWG on Safety recommended that ‘world-wide guidelines specify the rat as the rodent species to be used in the study for carcinogenic potential.’ (Emmerson, 1992, p. 208).

Such recommendations were inconsistent with the conclusions drawn by other expert scientists at that time and subsequently. Scales and Griffin (1992, p. 138) found that ‘based upon statistical evidence alone the use of only one rodent species for the assessment of the carcinogenic potential of non-genotoxic agents cannot be recommended’. In addition, in a survey of carcinogenicity studies of 22 pharmaceuticals in rats and mice, Bentley et al. (1992) reported at least 10 per cent of the true human carcinogens in the sample of rodent carcinogens were detected by mouse-only carcinogenicity findings.

Nevertheless, between 1992 and 1996, the ICH entrusted the US Pharmaceutical Manufacturers Association with the task of co-ordinating a systematic collection of evidence that could justify their conviction that the lifespan carcinogenicity test in mice could be dropped from regulatory requirements (Hayashi, 1994). Under the auspices of ICH, a retrospective survey of databases of pharmaceuticals tested in both mice and rats, and positive in at least one, was conducted. The databases consulted were those of the Japanese Pharmaceutical Manufacturers Association (JPMA), the FDA, the Dutch and German regulatory agencies and the CMR.7 The ICH used these databases to examine the extent to which: (1) mouse tumour findings gave rise to regulatory decisions against approval or marketing of pharmaceuticals; (2) mouse studies were needed to interpret positive rat studies; and (3) negative mouse studies were ignored in the face of positive rat tumour findings (Van Oosterhout et al., 1997, p. 8).

In their review of the Dutch and German regulatory databases, the ICH experts found that 13 out of 181 pharmaceuticals generated tumours in mice but not rats. Eight induced liver tumours, three induced lung tumours and three induced mammary tumours. For only one of these 13 compounds (8 per cent) did the positive mouse findings lead to regulatory action. In all other cases, the regulators judged that ‘the findings appeared to be species-specific, not occurring in rats, and therefore probably [ emphases added] not relevant to humans’ (Van Oosterhout et al., 1997, p. 10). Ten of the 41 pharmaceuticals (25 per cent) which generated tumours in rats but not mice led to regulatory withdrawal or restriction. This implies that positive carcinogenicity findings in rats have been of more regulatory significance than those in mice.

Regarding the impact of negative tumour site findings in mice, when there was a positive rat study, it was found that for 40 per cent (27 of 70) of these cases the regulatory assessments gave some weight to an argument that the absence of tumours in mice in the same organ indicated that the tumour findings observed in rats were probably not relevant to humans. The ICH experts interpreted this as a low percentage and consequently considered it to be further evidence that mouse studies are of little regulatory significance (Van Oosterhout et al., 1997). Regarding the 14 trans-species rodent carcinogens in the database, it was found that regulatory judgements were usually based on considera-

---

7 The Japanese industry and CMR reviews related only to industry knowledge of drug development prior to full regulatory review, so most of our discussion concerns the databases and regulatory positions of the European regulators and the FDA.
tions other than the fact that tumour sites in mice corresponded to those in rats, such as: mechanism of cancer induction assumed to be rodent-specific; oversensitivity of rodents to liver tumours; sufficient safety margin because tumours generated only in high doses in the rodents (Van Oosterhout et al., 1997, p. 15). Yet according to some expert toxicologists, ‘it would not be safe to assume that species barriers are never crossed or that there are invariably threshold levels below which non-genotoxic carcinogens pose no hazard for man’ (Roe, 1992, p. 105).

Nevertheless, ICH experts reported that the German and Dutch regulatory authorities did not pay any special attention to the fact that a compound increased tumour incidence in both species rather than one when assessing carcinogenic hazard to humans. Consequently, the overall conclusion reported to ICH regarding these Dutch and German regulatory databases was that ‘the relevance of the carcinogenicity study in mice was low’, and that one carcinogenicity study in the rat would be sufficient (Van Oosterhout et al., 1997, p. 16).

The results of the review by the JPMA and CMR of 178 pharmaceuticals in their databases indicated that positive tumour findings solely in mice had little effect on drugs’ development (Usui, Griffiths & Lumley, 1996, p. 280). This interpretation was based partly on the view that the majority of the ‘mouse-only’ tumours were found in the liver, and the belief that ‘rodent liver tumours have little or no relevance for the assessment of carcinogenic hazard to humans’. Thus, with the support of the Japanese regulators, they recommended that ICH should endorse a reduction in carcinogenicity testing from two rodent species to one, usually the rat (Usui et al. (1996), p. 283).

Representatives of the FDA also presented a retrospective analysis of their database of 125 pharmaceuticals. They reported that 78 per cent of all pharmaceuticals with positive findings in rat and mouse carcinogenicity studies would have been identified by a rat study alone, whereas 64 per cent would have been identified by a mouse study alone. Hence, the FDA’s review of its database for ICH implied that without the mouse study nearly a quarter of the carcinogenic effects from new drugs in their files would not have been detected (Contrera, Jacobs & DeGeorge, 1997). Furthermore, according to the FDA’s scientific experts at ICH, Contrera et al. (1997), a relatively high proportion of the trans-species carcinogens in the FDA database were either non-approved or approved with restricted clinical indications related to carcinogenicity findings. Evidently, FDA regulators paid special attention to trans-species carcinogens in risk assessment and noted ‘a major regulatory concern’ in relying on a single-rodent-species test because it would not be possible to identify transspecies carcinogens (Contrera et al., 1997, pp. 130 & 138–39).

FDA scientists were also influenced by tumour findings solely in mice and had used them to require manufacturers to conduct additional testing which indicated ‘that compounds were acting through a mechanism which may be relevant to humans’, or ‘even in recommending non-approval of certain agents’ (Ekman, 1996, p. 300). DeGeorge noted that he could not say in advance for any given pharmaceutical whether it would be more appropriate to require a single lifespan carcinogenicity study in rats instead of mice or vice versa, and insisted that ‘you can’t just automatically say “tumours are irrelevant because they are in the mouse or the rat”’ (Ekman, 1996, p. 299).

The accounts of these FDA representatives at ICH imply that the Dutch and German regulators seem to have been more permissive regarding trans-species rodent carcinogens because, unlike the FDA, they did not give them special attention in risk assessment, even though precautionary regulation would clearly have warranted such attention. Also the industry toxicologists and European regulators seem to have been much more willing than the FDA to dismiss liver tumours in mice as irrelevant to human risk. Whereas the FDA noted how tumours in mice, including some of the liver, were significant in such regulatory assessment. The European and Japanese regulators, who enjoy considerable flexibility in their regulatory judgements, safe in the knowledge that they are protected by secrecy legislation, may have been more accommodating than the FDA to industry requests to reduce carcinogenicity testing because, unlike in the US, regulatory decisions are much less likely to be subjected to public scrutiny by health advocacy groups via freedom of information laws or by legislative oversight and judicial review (Abraham & Lewis, 1998). The FDA, however, is inclined to be much more cautious and procedural because new regulatory policies may have to withstand challenges from consumer organisations in the courts or from Congress (Abraham, 1995; Abraham, & Sheppard, 1999).

In particular, the interpretation of mouse tumours by the Dutch and German regulators is suggestive, though not conclusive proof, of a permissive, rather than precautionary approach to carcinogenic risk assessment. Consider the fact that in only one out of the 13 (8 per cent) mouse-only carcinogens in the Dutch–German pharmaceutical database did the positive mouse finding count against the negative carcinogenicity results in rats and lead to regulatory action. By contrast, 27 of the 70 (40 per cent) of the negative tumour site findings in mice were utilised to count against positive carcinogenicity results in rats. In other words, this suggests that when a carcinogenicity study in mice was positive, but the rat study negative, the former was almost always discounted by the Dutch and German regulators, whereas when a carcinogenicity study in mice was negative, but the rat
study positive, quite often these regulators would entertain arguments which could undermine the positive rat study, by reference to the negative findings in the mouse. In both scenarios the effect is a regulatory assessment in favour of the drug because of a bias towards an overall negative carcinogenicity assessment.

This raises the possibility that the positive carcinogenicity findings solely in mice and the occurrence of trans-species rodent carcinogens have not received much attention from European regulators and industry toxicologists (in Japan and elsewhere) partly because of permissive industrial and regulatory practices, rather than solely because they are irrelevant to human carcinogenic risk assessment. From the perspective of protecting public health, therefore, there could be a fundamental flaw in basing new regulatory standards for future carcinogenicity testing on a retrospective analysis of regulators’ practices. The ICH may be creating doubly permissive drug regulation in the field of carcinogenic risk assessment by reformulating past permissive practices into the regulatory standards for the future.

Leaving these fundamental problems aside, the FDA’s position implied that the reduction from two rodent lifespan carcinogenicity studies to a single one in rats could not be done without losing important regulatory information which might compromise safety because testing in the mouse was sometimes, and unpredictably, relevant to human carcinogenic risk assessment. Hence, the FDA’s report to the ICH also threatened to undermine the entire ICH project to reduce lifespan carcinogenicity testing requirements from two to one study. Although DeGeorge (1996) was concerned that it carcinogenicity testing requirements from two to one undermines the entire ICH project to reduce lifespan carcinogenicity tests, by reference to the negative findings in the mouse. In both scenarios the effect is a regulatory assessment in favour of the drug because of a bias towards an overall negative carcinogenicity assessment.

This raises the possibility that the positive carcinogenicity findings solely in mice and the occurrence of trans-species rodent carcinogens have not received much attention from European regulators and industry toxicologists (in Japan and elsewhere) partly because of permissive industrial and regulatory practices, rather than solely because they are irrelevant to human carcinogenic risk assessment. From the perspective of protecting public health, therefore, there could be a fundamental flaw in basing new regulatory standards for future carcinogenicity testing on a retrospective analysis of regulators’ practices. The ICH may be creating doubly permissive drug regulation in the field of carcinogenic risk assessment by reformulating past permissive practices into the regulatory standards for the future.

Leaving these fundamental problems aside, the FDA’s position implied that the reduction from two rodent lifespan carcinogenicity studies to a single one in rats could not be done without losing important regulatory information which might compromise safety because testing in the mouse was sometimes, and unpredictably, relevant to human carcinogenic risk assessment. Hence, the FDA’s report to the ICH also threatened to undermine the entire ICH project to reduce lifespan carcinogenicity testing requirements from two to one study. Although DeGeorge (1996) was concerned that it would not be possible to identify trans-species carcinogens with a single-species carcinogenicity study, he conceded that this important regulatory information may not need to be derived from another lifespan rodent carcinogenicity study. FDA scientists at ICH accepted for regulatory purposes that a single rodent lifespan carcinogenicity study could be carried out combined with a new type of short-term in vivo carcinogenicity study in another rodent species (Contrera et al., 1997). According to a senior industrial scientist, who was closely involved with the ICH process, the industry representatives and European regulators were willing to drop the lifespan mouse study and introduce a new regulatory standard requiring nothing more than just one lifespan carcinogenicity study in rats. However, when the FDA objected to this, a ‘compromise’ was struck involving the new short-term carcinogenicity tests, which cost only about a fifth of the lifespan studies, and which many in industry hoped would be required only if the single lifespan test were positive.8

---

8 Interview with Vice-President of Clinical Safety, Janssen, 23 January 1998.

New genetics and alternative short-term animal studies

The short-term in vivo carcinogenicity tests in animals proposed as alternatives to the lifespan carcinogenicity studies involve genetic manipulation to produce transgenic rodents, which contain genes that are associated with tumour development, or which are deficient in genes thought to suppress tumour development (Tennent, 1996). Their rationale is that the early stages of tumour development, known as ‘initiation’ can be built into genetically engineered rodents so that carcinogenic effects can be detected in the whole live animal much sooner because only the later stages of carcinogenesis need to occur. On this logic, if a new drug is carcinogenic, then it should be detected fairly quickly by these short-term tests because the initiated animals should develop more cancer tumours more rapidly than the control animals (Schou, 1992).

In 1997 the ICH EWG on carcinogenicity testing reviewed the latest short-term in vivo carcinogenicity tests. One with rats, lasting 8 weeks, detected 26 of 31 (84 per cent) chemicals known to be non-genotoxic liver carcinogens in rats from lifespan studies; a second, involving rats for 36 weeks, detected only 4 of 7 (57 per cent) of the non-genotoxic chemicals known to be carcinogenic across a number of different organs from lifespan studies in rats; a third, using transgenic mice for six months detected only 4 of 5 (80 per cent) of the non-genotoxic carcinogens known from lifespan studies in mice; and two others involving mice could not detect non-genotoxic carcinogens (Mitsumori, 1998). The ICH concluded that the short-term tests in rats had ‘not been fully validated’ and that ‘their utilisation’ was ‘problematic’; while the transgenic mouse tests required further validation before any conclusions could be drawn about their usefulness (MacDonald, 1998; Mitsumori, 1998, pp. 266–67). As Weissinger (1992) explains, validation should include studies conducted to show that tumours previously noted near the end of a lifespan study are not likely to be missed in chronic carcinogenicity studies over a shorter period. As recently as 2000, they had not been validated (IFPMA, 2000, p. 7).

The drawbacks of these new short-term carcinogenicity tests may go much deeper than non-detection rates of non-genotoxic carcinogens. According to Huff, Haseman and Rall (1991), short-term studies in animals are inferior to life-span studies because ‘initiation’ is accomplished artificially and the chronic exposure, which best mimics the human situation, is not possible. Robertson (1992, p. 230), Director of Drug Safety at Merck, Sharp & Dohme, has also noted that many tumours identified in studies with pharmaceuticals are produced very late in the studies, implying that the interaction of exposure and ageing is important. It has been suggested by Roe (1992) that this may be because defence mechanisms which are effective earlier in life...
cease to be so later on. Consequently, he favours lifespan tests designed to assess whether a drug weakens or strengthens these defence mechanisms by collecting data which relate ageing to carcinogenicity. For Dayan (1992), the lifespan test is further justified by the time required for weaker carcinogens to produce an identifiable effect.

Despite the many concerns about the appropriateness, efficacy and validity of the new short term tests, in July 1997, the ICH approved the guideline, ‘Testing for Carcinogenicity of Pharmaceuticals’, which proposes that in place of a second lifespan carcinogenicity study, a short term study in a rodent model may be appropriate (Lumley & Van Cauteren, 1997).

Discussion and conclusion

The ICH management of the regulatory standards for carcinogenicity testing have been about reducing testing requirements, rather than the harmonisation of inconsistent standards across regions. Indeed, one could argue that ICH has produced less ‘harmony’ than existed before because of the difference of opinion between FDA scientists and European regulators about the value of carcinogenicity studies in mice. The supporters of ICH acknowledge that it aims to make drug development and regulation more cost-efficient, but they also claim that it does this without compromising safety standards. However, the criticisms made by ICH experts of the ‘conventional’ rodent lifespan carcinogenicity tests are by no means compelling and highly disputed.

Three mutually supporting reasons suggest that the ICH’s approach to the lifespan carcinogenicity test in mice may reflect permissive regulation leading to a diminution in safety standards. Firstly, the ICH analysis is based on accounts of past practice by industry and regulators and there is evidence that this practice has been permissive with respect to the regulatory standards of the time. Secondly, the ICH analysis itself reveals that European regulators have been much more permissive in regulating trans-species rodent carcinogens than the FDA. And thirdly, a five-fold higher percentage of negative carcinogenicity findings in mice have been used by European regulators to suggest that positive carcinogenicity findings in rats are irrelevant to human safety, than have positive carcinogenicity findings in mice been used to suggest that negative carcinogenicity findings in rats may be wrong.

Consequently, there is reason to believe that ICH may be converting permissive regulation into new scientific standards for carcinogenic risk assessment. It is difficult to see how this will not lower safety standards unless the new short-term tests prove to be breakthroughs so far not forthcoming. Significantly, however, the ICH guideline to reduce lifespan carcinogenicity testing was accepted by regulators in the EU, US and Japan by 1998—before validation of the new short-term tests was completed.

Regarding the central hypotheses being tested in this paper, therefore, we conclude that, in the field of carcinogenicity testing, the ICH management of international harmonisation of medicines regulation is not achieving the simultaneous improvements in safety standards and acceleration of drug development hoped for by Davis (1997). Rather, our research supports the more sceptical view that the latter is being achieved at the expense of the former. Further research is needed before generalising our findings to the ICH process as a whole. Nevertheless, given the enormous impact of ICH guidelines on medicines regulation world-wide, our limited findings are significant, especially as many expert scientific advisers to drug regulatory agencies seem to have accepted uncritically the conclusions reached at ICH. Moreover, there is evidence of a similar deteriorating trend in other areas of ICH’s safety standards, such as chronic toxicity testing in non-rodents, exposure of patients to carcinogenic risk in clinical trials, minimum duration of patient treatment in clinical trials and requirements for reporting of adverse drug reactions (Abraham & Reed, 2001).

In conclusion, our research suggests that neo-liberal influences on the regulatory agencies, such as the goal of increasing and accelerating access to markets for industry, are neither neutral nor coincidental in their effects on regulatory science. Rather, such ideological influences are associated with a lowering of regulatory standards in the area of safety testing. During approximately the last 15 years, this ideology has penetrated the scientific knowledge base of carcinogenic risk assessment itself by first creating a political environment in which a permissive gap between the practices of regulatory scientists and the cognitive norms of their regulatory science has become administratively acceptable, especially in Europe; and then introducing a new world-wide regulatory science in which the cognitive norms have been made consistent with those permissive practices.

Administratively, it seems likely that regulators have been persuaded to adopt these changes in regulatory science because they are convinced that they must operate within a neo-liberal political environment for the foreseeable future making them vulnerable to industry demands to accelerate drug approvals. Reductions in drug testing requirements and increased flexibility in interpretations of what counts as evidence of carcinogenic risk make it easier for regulators to approve drugs—and approve them more quickly. A related explanatory factor is that the ICH process was dominated by industry-regulatory dialogue, and excluded consumer organisations and public health advocacy groups. This is likely to have strengthened the industrial agenda of accelerated approval in the
minds of regulators due to an absence of alternative political and public health perspectives. For example, whatever one’s view of the reductions in requirements for rodent lifespan carcinogenicity testing, the ICH contained no discussions of whether there should be more stringent post-marketing surveillance standards for the detection of carcinogenicity.

Rather than accepting problematic new methods of carcinogenic risk assessment prior to validation, regulatory agencies should involve a full range of expertise in medicine, pharmacology, toxicology and public health advocacy, as well as industry in order to evaluate which methods are in the best interests of public health. This would reduce the likelihood that developments in the regulatory science of carcinogenic risk assessment would be shaped by neo-liberal influences and trade interests at the expense of patient safety. Furthermore, the regulatory agencies should use their retrospective analyses to identify those drugs on the market which were carcinogenic to animals, and which offer no clinical risk-benefit advantage over alternative therapies available to patients. Within a precautionary regulatory strategy such drugs could be withdrawn from the market or denied marketing approval.

Acknowledgements

Thanks to the UK Economic and Social Research Council (Grant ref L323253020) for funding some research for this paper, and to Peter Davis and two anonymous referees for comments on an earlier version.

References


