Toxicology: Statistical perspectives

Pranab Kumar Sen
Departments of Biostatistics and Statistics, University of North Carolina, Chapel Hill, NC27599-7400, USA

Toxicity abounds in nature, environment, and in our modern life-style. Toxicology relates to the study of the intake process of such toxins by human being, their mode of propagation, biological reactions, molecular level of penetration, genotoxicity and aftermaths. Because of the latent nature of a large class of toxic substances, the extreme variability of human metabolism as well as their exposure to toxic material, yet unknown nature of many carcinogenic activities, and immense difficulties in the assessment of effective toxicity levels (especially in the environment), there is a genuine need to have statistical appraisal at each phase. These statistical perspectives are highlighted here along with some outline of recent statistical approaches in this much needed assessment task.

FROM time immemorial, toxins have been recognized by the human being, and used both beneficially and destructively. Like morphine and other pain relievers, such toxins have often been used in medicine and allied fields to combat some diseases or disorders, and more often in punitive or destructive modes; the use of (slow as well as instantaneous) poisons to eliminate an undesired person has been in practice from the dawn of human civilization. Unknowingly, sometimes we might be in touch with some toxic plants, fruits or material, and such contacts might have widespread and often disastrous effects. Even in the animal world, toxins are recognized by various species who might have acquired such knowledge through their ancestors. From the ancient Vedic time, toxins in various herbs and plants have been most thoroughly studied by Indian herbal physicians (known as the *Vaidyas*) and their acquired knowledge (*Ayurveda*) is still a yardstick for the (western as well as oriental) medical system; the basic difference between the two lies in the occidental use of toxic chemical compounds (inorganic toxins) instead of the organic toxins in herb and plants. The entire foundation of the homeopathic medicinal system is based on the impact of toxins on human body as well as mind, and is a classical example, how a form of a toxic element can be used to nullify some other toxicity in our body. Although there may be some controversy over the doctrine of homeopathic medicine system, there is no doubt about the neutralizing capacities of toxins of different kinds.

For correspondence. (e-mail: pksen@bios.unc.edu)

Toxins are characterized by their toxicity, in some form or the other, in relation to their reactions on living organisms including man. In a traditional sense, toxicity relates to the poisonous effects which are associated with specific chemical compounds (such as potassium cyanide), various plants and fungi, and with our environment and ecosystem. A good deal of knowledge has been acquired on toxicity of various chemicals, fungi and other compounds. Such toxicities are fairly quick in progression and fairly deterministic in the dose-response relation. As such, viewed from statistical perspectives, such quick-action, fairly deterministic toxicity are not of much interest. Rather, the more complex types that often progress slowly and invisibly, and in that process invite other activities, such as genotoxicity and carcinogeneity, have a far greater need for statistical appraisal, and we shall mainly confine ourselves to this aspect of toxicology. This complex of toxicity includes a wide variety of subdisciplines: *Plant toxicology* and *animal toxicology* refer to toxins and their toxicity that are prevalent in plants and animals respectively; taken together, they constitute the *organic toxicity*. Toxicity is also prevalent to varying degrees in drugs and pharmaceutical products. Toxicity may arise in the use of pest control and storage facilities of agricultural products; it can also arise in the process of cooking food, its preservation and service. Industrial factories and plants are notorious sources of toxicity; industrial waste also often lead to significant amount of toxicity. More significantly, *environmental toxicity*, resulting from air pollution or toxicants (such as industrial and automobile exhausts, environmental smoking, thinning of ozone layer, smog containing airborne particulate matters, acid rain and others), and groundwater and subsoil contamination (due to arsenic minerals, industrial waste dumping, landfill practice, ecological disasters, nuclear waste disposal, and other sources of associated toxicants), might be labelled as a significant ingredient of toxicity (though such toxicity effects are generally slow in progression and hard to isolate in detection). In the foothills of the Himalayas and the Gangetic valleys, particularly in the lower basins, arsenic contamination of groundwater and soil from subterranean source is a growing threat to public health. Besides, organic arsenic occurs in plants, fish, crab, human body and other organisms, though this may be generally less toxic in effect compared to arsenic minerals. Arsenic contamination of groundwater may be caused by geological and anthropometric processes. This contamination is due to a chemical process wherein arseneous acids from buried

(mostly ferric) arsenate are produced. This is indeed a major concern in a greater part of (eastern) India. Furthermore, toxicity in drugs and other intoxicating beverage materials (particularly, illicit liquors) is also an important factor. Occupational environment may have a significant contribution toward this toxicity. In natural disasters and calamities (like earthquakes, landslides and hurricanes) *aqua-toxicity* may result from contamination of animal corpse and wastes, human wastes as well as ecological dysfunctions. Industrial waste containing toxic substances and decomposed organic products may find their way either directly by the flow of surface water or through subsoil layers (in dumping) to water sites; such toxicity may seriously affect the marine ecology (biology), and through that human health and quality of life as well. Arsenic contamination is a serious problem in this respect. Our modern life-style has been inducing some forms of toxicity through our addictions to modern amenities resulting from excessive use of chemicals, electronics and mechanizations. The level of toxicity and its intensity may vary considerably from one source to another, and their impact potential may also differ considerably. The intake process of toxicity by human being (and other living objects) may also vary, and we shall refer to that later on. The modes of intake of toxicity by human being may follow the *inhalation*, *absorption*, or *ingestion* tracks, and often in a combination of more than one mode at a time, resulting in *synergism* which may complicate the picture considerably.

Toxicology deals with the study of the complex of resources, synergism, their distribution, intake processes, uptake sites, metabolic reactions, and aftermaths of toxins around us. Often we are susceptible to them, though unnoticeably, by our living environment, and often we are addicted to them by our life-style and other irrational convictions. Within this broad discipline of toxicology, we have a battery of allied fields: *behavioural toxicology*, *genetic toxicology*, *plant toxicology*, *neurotoxicology*, and *reproductive toxicology*, among others. Though the genesis of toxicology is in the biochemical sciences, its traditional frontiers have been greatly expanded and fortified by the addition of new territories: *neurotoxicity* and *genotoxicity*. The alarmingly escalating level of toxicity of our bioenvironment is a life-threatening concern to the entire mankind and our ecosystem as a whole. Both the disciplines of genotoxicity and neurotoxicity are devoted to certain aspects of toxic effects on human body and mind that go far beyond the conventional biochemical levels. With the advent of computer intensive modern molecular biology and genetics, more research developments in this fertile area are bound to take place, and *bioinformatics* may have the triumph card in this respect. Statistical reasoning plays a key role in this interdisciplinary field, though it needs to be properly blended with the basic understanding of (inorganic and organic) toxicology, human as well as subhuman and plant physiology, neurobiology, and modern information technology. Even *epidemiology* and *environmental health sciences* have good standing in this venture, and statistical reasoning must adhere to these disciplines as well. As such, we like to dwell deep into statistical perspectives in the light of the broad interdisciplinary aspects of toxicology.

The toxicity uptake passage

The propagation of toxicity intake is different for different species; what may be toxic for human being may not be so for some other organisms and vice versa. Even among the subhuman primates, there is a qualitative and quantitative difference in toxicology, and as a result, it might not be prudent always to draw conclusions from rodents and apply them to human beings. Let us focus our attention here on human exposure to toxins and toxicology for human being. In doing so, we may not be able to bypass totally animal perspectives, and that will be discussed later on. Toxicity is perceived by human exposures to toxic elements, and the toxicity intake process may take one of the following three routes: (i) Skin absorption (mostly epidermal contact); (ii) Inhalation (mostly, through respiratory system); (iii) Ingestion (mostly, through digestive system).

In the absorption process, toxic elements make their way through the skin, contact the minute blood vessels, and in that process reach the cardiovascular system with access to liver, kidney, brain, developing organs, and mammary glands (in women). In the inhalation mode (which is believed to be the most significant one), toxic elements inhaled (mostly) through nose (and mouth) make their way through the nostrils all the way to the bronchioles along the respiratory tract. As the bronchioles lead to millions of alveoli (air sacks) of the lung tissue, such toxic elements may block the passage of exchange of inhaled oxygen with the exhaled carbon dioxide and vapour; with reduced supply of oxygen and presence of toxic elements, the contaminated blood stream can trigger off all consequential problems. The formation of tumours (leading to lung cancer) is also supposed to be linked to such less-than-ideal activity of the alveoli. Through the blood vessels, the inhaled toxic elements can reach other organs and the aftermath can continue. In the ingestion mode, toxic elements enter into human bodies mostly through the mouth and digestive system; their target is the gastrointestinal (GI) tract which has access to blood stream, so that toxicity can again prosper through blood circulation. Many common forms of toxicity, such as food-poisoning and toxic chemicals in food and drinks, follow the ingestion track, and may often be quite fast in their progression. In environmental toxicity, on the contrary, the progression is slow and often unnoticed, and it is hardly the case that a single mode of progression is followed, and it may even be difficult to sort out the three

specific components in any typical situation. For example, in arsenic contamination of groundwater, ingestion is prominent in the drinking mode, while absorption takes place from external use of contaminated water, such as bathing, washing faces, etc. In drug-addiction problems, we might have more distinct and identifiable intake process of the toxic elements. For example, in *moonshinning* (or illicit liquor drinking), the intoxicating alcoholic drink intake is directly through the mouth to the stomach and liver areas, and as a result, ingestion is the primary track. In case of smoking of tobacco products, the tract is primarily the respiratory channel, so that inhalation mode is the dominant one. Chewing tobacco products, as is the practice among certain athletes, may however, have a significant ingestion mode of intake. Outdoor exposure to sun, especially when the ozone layer is thin, the ultraviolet radiation hits the body mostly through the exposed part (skin and hair), so that absorption mode is more prevalent than the others. In intravenous intake of narcotic drugs, the toxicity may directly hit the blood vessels bypassing all the three common uptake modes mentioned before. Toxicity in drugs (used for medicinal purpose, not for addiction) that are mostly taken orally, follows generally the ingestion mode, so does the toxicity in food processing, cooking and eating practices. On the other hand, in usual atmospheric pollution, we have a more complex picture, and we present an overview of that first. At the present time, our primary concern centres around the complex of environmental and behavioural toxicity.

Environmental toxicity

While we might have some control over the use of common chemicals for our household work, safer drugs for our medication, natural foods for our nutrition, which may result in reduced and relatively safer levels of toxicity, we may not have that much control over the environmental toxicity which is escalating at an alarming rate due to our insufficient efforts to keep our bioenvironment clean, and also due to our indifference to our mother planet, the Earth. Let me illustrate some of the important factors in the light of the intake processes described earlier.

Environmental smoking

Smoking of cigarette, cigars, bidis and allied tobacco products leads to significant toxicants, not only for the active smokers but also for passive smokers (that is, those who are in the vicinity of active smokers); the mode of intake is primarily inhalation of toxic fumes and carbon monoxide-rich gases. Emission from factories that use coal or wood as fuel, as well as home cooking on coal energy also contribute significantly to environmental smoking. It is a belief that cigarette smoking is linked to lung cancer and other respiratory diseases, though the cancer etiology is still not that precisely known. The Surgeon General in USA has a statutory label on every cigarette packet: Smoking is dangerous for your health. Are we aware that industrial smoking (for example, fumes from factory exhausts) may be equally, if not more, harmful to our health?

Automobile and aviation exhausts

A majority of automobiles (including cars, vans and trucks), and locomotives (including trains) run on diesel and gasoline fuels; so do the aeroplanes of various types (which generally require a high-octane fuel product). Only in some countries, alcohol made from sugarcanes is used as fuel in cars. These conventional fuels contain various toxicants. On top of that in severe winter time, some other reagents are added to gasoline products to facilitate smooth starting in sub-zero temperature; these are also suspected to have carcinogens and other toxicants. Elimination of lead from gasoline products (on a wider scale) has resulted in some improvements in recent past. Airplane exhausts, while on air, contribute toxicants in the upper atmosphere that with rain and smog try to come down towards the surface level.

Occupational environment and toxicity

Many of the present time occupational environments are contaminated with toxicants and other health hazards. As examples, we may cite the following:

Coal miners (*black-lung*) *disease*: In mining operations all over the world, and more so in third world countries, in the underground tunnels, there may be an excessive amount of airborne carbon particulate materials (APM), carbon monoxide, and other harmful gases. In that way, inhalation toxicology is most pertinent to the coal miners occupational environment, and lung diseases used to be prevalent among them. Subsoil water contamination and moisture, coupled with possibly inadequate air circulation system may add more distress to this toxic environment.

Asbestosis: Asbestos is an incombustible, fibrous mineral that used to be commonly used as an insulation material, and often as a substitute for tin or other metals for corrugated roofing. The white dust particles in an asbestos plant are quite catalytic for respiratory diseases as they may get absorbed in the lung tissues. Again, inhalation mode is the more common intake process, and asbestos has also been identified as a carcinogen. Even in houses or schools and work places where asbestos used to be the insulation material, the dust particles play the same role, but possibly to a lesser intensive level. Young children often unknowingly placing their hands on such insulations and without washing taking food may induce the ingestion

mode, while epidermal contact may have absorption intake mode.

Chemical plants for agricultural fertilizers and pest control products: DDT is a classic example of such a highly toxic material. Though primarily meant for controlling insects and bugs, it has a significant carcinogenic effect on human beings as well. The list is by no means very tenuous. The 1984 highly toxic gas leakage from the Bhopal Union Carbide plant had such a high toll in terms of mortality and morbidity that the deleterious effects of such toxic gases are evident.

Nuclear power plants: No matter how much precaution is adopted in dumping the nuclear waste products or releasing the (low level) radioactive gases in the atmosphere, such radiation effects can affect human being in inhalation, absorption and ingestion modes. The casualty from various power plant leakage bears this testimony. The impact of nuclear warheads and weapons (including bombs) can be even more disastrous.

There are literally thousands of occupational hazards and toxicity sources, and it might be better to check occupational safety and security standards to minimize their impact.

Waste products: Industrial and domestic waste products generally produce a lot of toxic effects. Sulphur PMs, nitrogen dioxide, mercury, carbon monoxide from oil refineries as well as thermal power plants, nuclear and chemical plants waste products, and even the garbage collected from households contain significant toxic materials. All modes of intake are active in this respect. We refer to (refs 1 and 2) for some discussion.

Behavioural, reproductive and neuro-toxicology

Our engulfing toxic environment often works silently in slow progression at an early phase of intervention, and then at an opportune time, switches its gear and pounds into deleterious impacts. The intimate relationship between body and mind in the entire animal world (including human being) has been well established. A healthy body (especially at the developmental stage) tends to have harmonious impact on a healthy mind. Continual health problems, particularly relating to chronic diseases and irreversible disorders, may significantly affect the quality of life and behavioural patterns. As much as malnutrition has a significant influence on mental health-related illness, toxicity has a similar impact on personality, outlook on life, and quality of life; this branch of toxicology is referred to as the behavioural toxicology. Domestic violence, irrational thinking processes (including committing suicide), clinical depression, addiction to drugs and alcoholic drinks may all be linked to the broad spectrum of toxicity that surrounds us in our everyday living. In fact, the aging process is also tied-down to the environmental toxicity in a broad sense, though a direct causal relationship may be hard to establish; genetic toxicology is coming up with good evidence of such behavioural dysfunctions due to toxicity in our bioenvironment. Epidemiology has long been involved in this field of research through mostly observational or case-control studies. Environmental and genetic epidemiology are also going through evolutionaly growth in acquiring fruitful information in this respect.

Reproductive (and developmental) toxicology deal with adverse effects on male and female resulting from exposures to specific toxicants that may lead to reduced fertility, alterations in sexual behaviour, adverse pregnancy outcomes and alterations of other functions relating to the reproductive system (including depression)[3]. On the other hand, developmental toxicity relates to adverse effects in the developing organism due to exposures to either parent before the actual conception or during the prenatal or the postnatal development; these include still-birth, birth defects, teratogenesis and structural abnormality, altered growth, and abnormal functions. For people exposed regularly to anaesthetic gases, a decrease in fertility and premature birth problems have been observed; among the smokers in their pregnancy, early pregnancy loss, reduced body weight for babies and other birth defects are more commonly observed. Toxic gases, such as dioxin, may have adverse reproductive and developmental toxicity effects.

Neurotoxicology deals with the impact of toxicity (mostly, prevailing in surrounding environment) on human nervous system which is controlled by the central nervous system (CNS) located in the proximity of the brain. Toxic elements in the cardiovascular system resulting from either inhalation of toxic gases or PMs, or ingestion of toxic foods or drinks, and also absorption of toxicity through skin, may generally impair the normal working environment of the CNS. Reduced flow of oxygen to the cortex area is a significant factor in this respect, and compounded with other toxic elements in the blood in the upper aorta, it can have even far more devastating effects. A less-than-ideal level of functioning of the CNS and the cortex can also affect the personality and sense of security. Thus, behavioural toxicity and neurotoxicity are quite interrelated. Even reproductive toxicity has a good bearing on both behavioural and neurotoxicity.

Genetic toxicology

The impact of environmental toxicity, in a broad sense (as depicted earlier), has crossed the traditional biochemical and biomedical boundaries and permeated well into the frontiers of genotoxicity. The very basic fact that often toxicity interacts with the genetic material and transmits it

in a chain of cell divisions with potential carry over effects on the future generations as well. Moreover, such toxicity studies at the molecular level may sometimes provide clues for cancer etiology, and modern molecular biology has indeed come up with a challenging task to incorporate genetic toxicology in deeper understanding of progression of genotoxicity in living organisms.

The advent of molecular biology and cytogenetics has opened up a wide area of research potential for understanding how far toxicity can be related to the genes, the carriers of hereditary traits. On one hand, the present generation of scientists are running after the gene therapy wherein either a gene carrying a virus or a disease code can be repaired or better be replaced by a host of genes who are free from this virus or disease code. On the other hand, back in the mind of these super-scientists is the aspiration of building a super human race by such gene transfers or repairs. Obviously, even now, there are formidable roadblocks to transfering their dreams to reality. These new genes need to be transplanted in the cells containing the old ones, and in a predictable setup, these intruder genes should be able to eliminate the host ones. As this replacement process generally involves a gene-war against the host genes, it is quite possible that the immune system may turn to be rather aggressive, destroying not only the harmless virus in the intruder genes but also its own vital organs. Chemotherapy, often used for cancer treatment, has this characteristic kick-back. Gene therapy at the present time is also confronted with the same dilemma: preserving the vital organs in this gene-war while allowing the intruder genes to take over the control in a way so as to eliminate the target virus completely from a target organ. For this reason, genetic toxicology is more commonly investigated with subhuman primates, like rodents, monkeys and cats or dogs, along with the hope that the acquired findings would remain pertinent to human beings as well (a gene-way from mice to man!).

Genetic toxicology or mutagenesis focuses on the study of agents that damage DNA (deoxinucleic acid) and related genetic material. Such agents have the capability to alter the human gene pool with unknown but potentially deleterious consequences for future generations. There is a broad range of toxicants that lead to such DNA damages and alterations. In this context, there is a firm belief that inhalation is the primary industrial/environmental site of uptake of toxins. For that reason, usually genetic toxicology assays are related to inhalation toxicology: radon inhalation, inhalation of tobacco-related chemicals, formaldehyde inhalation, inhaled arsenic products, butadiene, benzene and mineral fibres. Also these assays were mostly based on animal exposures. Effective dose levels and exposure time periods are very crucial in such studies. There are lots of statistical issues related to the validity of such assays and their ability to transfer the findings to human exposures that could be far different from the laboratory animals. We shall discuss these briefly later on.

For human exposures, initially, genetic toxicology assays were employed to screen for environmental mutagens and presumptive carcinogens. DNA damage, however, may also be a factor in other diseases, such as aging, depression, etc. The use of biomarkers has made it possible to formulate genetic toxicology assays for human being in an epidemiologic setup and under some controlled experimental setups. Of course, as these assays are meant for human subjects, generally low dose studies are contemplated, and the basic issue of utilizing statistical information from such low dose assays to normal exposure ones remain a challenging task. In this context, we may also refer to bioassays that may be considered for *bioavailability* and *bioequivalence* studies, though there are some basic differences in the two setups. In clinical therapeutic equivalence trials, bioassays mainly relate to assessing relative potency, while in bioequivalence trials, usually, the relative bioavailability of different formulations of a drug are compared, so that *pharmacologic* results of administrating essentially a common drug in alternative forms, such as capsules vs tablets or liquids, and/or dose levels and/or frequency of prescription capture the main interest. Bioequivalence trials are more commonly applied on human subjects with proper safeguards from initial animal and therapeutical studies[4].

Radioimmunoassays are labelled with (usually small) doses of radioisotopes; in *immunoradiometric* assays antibodies are labelled. For a broad range of antigens, such radioligand assays enable the estimation of potency from a very small quantity of material and usually with high precision. Such assays are based on radiation counts at various doses in a fixed time period. In this way, potency estimation involves the relationship between counts of radioactivity and dose, usually at low levels[5].

Assays for genotoxicity can be used as predictors or biomarkers of disease, especially cancer[6,7]. Although the relationship of biological markers to exposure and disease can be conceived in a simple model allowing susceptibility and environment life-style to be related to internal dose, biologically effective dose to biological responses, altered structure/function, to disease to prognosis – in reality, there are considerable complexities due to the multitude of interacting factors, the diversity of potential biomarkers, each with its own sensitivity and specificity for disease predictivity, and the difficulties in constructing prediction batteries of biomarkers. Genetic toxicology, in general, is being heavily influenced by the molecular revolution. DNA is made of two strands of nucleotides with a large number of positions or sites, each position on the strand containing four possible bases (labelled A, C, G and T). Thus, DNA represents sequences of qualitative responses, For such multiposition strands, it might not be wise to introduce simple first-order Markov models to model DNA sequence data, and hidden Markov models (HMM) have been advocated by a number of researchers[8]. There are certain conceptual gaps

in the statistical adoption of HMMs, and we shall discuss these later on.

Statistical perspectives

Given the above outline of toxicology in relation to our body, mind, our bioecosystem, and society as a whole, the basic concern is to identify the toxicity undercurrents, to have a complete inventory of toxins of various kinds, to arouse our awareness to the grave impact of such toxicants, to study their mode of progression and distribution [i.e. pharmacodynamic-pharmacokinetic (PDPK) aspects], their (molecular) biological interaction and impact, to quantify the risk due to toxicants, and effectively accomplish the risk assessment task with a view to promoting global safety standards and improving the quality of life. In all these aspects, statistical reasoning plays a fundamental role. Several broad aspects of statistical perspectives need to be appraised in this setup. Broadly speaking, collection of relevant toxicologic data that can be put to statistical modelling and analysis schemes needs careful planning as well as skillful sampling techniques (as a complete census is often ruled out). In many situations, we may have, at best, some observational studies (mostly posed from epidemiologic considerations), and it could be a challenging task to validate statistical analysis in order to procure objective conclusions. For this reason, epidemiology and biostatistics need to be blended to accomplish this task as effectively as possible. Animal or dosimetric studies are often advocated to gather information of therapeutic and toxic factors and their impact on the end-point(s) in mind. There are, however, some roadblocks to such dosimetric studies, especially in environmental toxicology, and hence, we need to appraise them too. Bioassays are more often used in pharmacologic setups, and are quite useful in toxicologic studies as well. There are some other statistical issues relating to toxicology that deserve some appraisal too. We present an outline of this statistical complex below.

Design and analysis of toxicological studies

Toxicologic studies are either made on animals (*in vivo* or *in vitro*) or on human subjects; the latter being under various observational setups rather than in strict experimental ones. Typically, such studies are made first on subhuman primates to set appropriate dose and exposure levels as well as examine possible side-effects or other forms of toxicity (besides the main end-point under consideration). These animal studies are generally referred to as *dosimetric assays*. Even such dosimetric studies differ considerably from conventional laboratory experiments which can be usually conducted under relatively more controlled experimental setups. Toxicity in environment and in other natural forms often have latent effects, usually working in a rather synergism mode, and often, may have difficulties in identifying the principal end-point(s). Because of these reasons, conventional designs that are used in agronomic or industrial studies or even in clinical/medical studies are often inappropriate for such toxicologic investigations. Although *longitudinal studies* or *repeated measurement designs* have more relevance to toxicologic investigations, simple statistical tools are usually not so appropriate in such contexts due to model complexities, considerable degree of spatial and temporal variations, and the very nature of the response pattern. We may conceive of a simple dose-response model wherein a simple linear or specific nonlinear regression function with distributional homogeneity of the residuals from the regression is conceived, and often even the error distribution is taken for granted to be Gaussian. This stationarity of the errors in a statistical sense is a fundamental assumption of statistical modelling and analysis schemes, and often blocked designs are used to reduce further the margin of errors[9]. In a multivariate setup[10], multivariate analysis of variance (MANOVA) models have been developed even under more complex designs relating to multi-factor and multi-response experiments. Here also, distributional homogeneity of the error vectors, additivity of the treatment and factor effects, and normality are taken for granted. In recent years, more emphasis has been laid down on robustness aspects that introduce flexibility to the models[11]. However, additivity and some other homogeneity conditions remain as vital components of the basic regularity assumptions. In toxicologic studies, especially relating to environmental toxicants, the degree of homogeneity of the distribution of errors, the additivity of the effects, and above all, Gaussian distributional patterns may all be questionable to a larger extent. As such, whatever optimality considerations pertaining to standard statistical designs may no longer be pertinent to toxicologic studies. This serious aspect deserves a careful appraisal, and will be made later.

Sampling schemes

Accurate determination of prevailing toxicant levels and their interactive mode of action need a careful statistical appraisal. In agronomic studies, laboratory experimentation, or in many other fields of application, simple random sampling (SRS), with or without replacement, is generally adopted. This permits standard statistical tools to draw statistical conclusions subject to assertable levels of margin of errors. In more complex studies, instead of SRS, one may have unequal probability sampling (UPS) schemes, again with or without replacement. In toxicologic studies, particularly in environmental toxicology, the prevalence of toxic elements may not follow a regular or even smooth pattern. Moreover, for all practical pur-

poses, it might not be feasible to record the prevalence data cloud on a continuous basis over a spatial and temporal domain. Usually, a certain number of grid-points are chosen as monitoring sites, and data collected from such points, spanning a time period, are incorporated in the study of spatio-temporal variation over the area spanned by the chosen grid-points. In view of the paucity of such grid-points, enormous (and uncontrolled) variability in the level of contaminants, and rather complex PDPK undercurrents, it is quite important to pay adequate attention to the sampling scheme and the number of observations that may be needed to draw statistical conclusions with some postulated margin of error. Further, in atmospheric toxicity, usually a three dimensional grid set is used to study the effect of altitude and the distance from a contaminating source, and in addition, nonstandard recording device may be needed. This may complicate the sampling scheme. In many cases a systematic sampling scheme is used for administrative convenience, though that might lead to inappropriateness of standard statistical methodology for drawing objective conclusions from acquired data sets.

Epidemiologic undercurrents

In a conventional setup, observational studies are usually made for toxicologic investigations involving human beings, while laboratory experiments are advocated for animal studies. Even for such dosimetric studies, a complete randomization may not be always adoptable, and intracluster dependence pattern may cause serious problems in statistical modelling and analysis. For epidemiologic studies, often, case-control, matching, and cohort analysis methodologies are recommended for better control of bias and undesirable variation. For purely observational studies, there may be a basic issue: how far statistical modelling and analysis can be incorporated in such studies? In view of the paramount importance of such scientific findings and their impact on our community as well as public health, there is a genuine need to explore appropriate and valid statistical modelling and analysis schemes that would match the type of observational studies employed; we refer to a recent volume[3] where such statistical issues have been addressed thoroughly (by a host of contributing authors). At present, there is considerable emphasis on genetic epidemiology and environmental epidemiology which have also good relevance to toxicology. (Bio-)statistics and epidemiology have been working in close cooperation in resolving some of the basic issues in such epidemiologic studies. One of the most significant developments in this context is the evolution of clinical epidemiology that borrows strength from pilot dosimetric studies and applies them to human subjects through controlled clinical trials. In this manner, there has been a harmonious blending of epidemiologic, clinical and statistical reasonings.

Biomechanistic and PBPK modelling

Advances in biotechnology and modern computer technology have made it possible to advocate suitable models that incorporate the basic biological and toxicologic factors in the depiction of the intake process of toxicants, their pathology and biological/toxicologic impact in terms of appropriate biomathematical formulations, and on top of that, impose stochastics to reflect the unaccounted variations associated with the process. These constitute the *physiologically-based pharmacokinetic* (PBPK) models. In the presence of genotoxicity, such models may generally become more complex due to linkage and other genetic stochastics factors that involve more dominant stochastic elements. For example, growth of a tumour possibly leading to cancer can be related to the intake of certain specific toxicants in certain specific mode, and such stochastic biomechanistic models may be quite useful, particularly for studies involving rodents or other subhuman primates. In lung diseases, such a biomechanistic model may become more complex due to genetic effects and highly variable environmental effects. Occupational hazards also contribute more toward a complex stochastic environment for which simple models may not work out properly. For this reason, PBPK modelling is often advocated to incorporate more biological and toxicologic understanding of the whole intake process and its aftermath. At present, there is a lot of emphasis on the so-called *stochastic partial differential equations* (SPDE) which have the natural ability to incorporate stochastics in nice diffusion process formulations, and at the same time, the deterministic factors in terms of the parameters that appear in such SPDEs. On the other hand, in a majority of cases, especially involving human subjects, such SPDEs are quite complex and involve a large number of parameters; statistical interpretations and drawing conclusion from such SPDE models may often become difficult. There is ample room for more developments in this much needed area of statistical research.

Dosimetric studies

Animal studies or dosimetry can be regarded as the precursors of serious scientific toxicologic studies. Basically, most of the toxicity studies, though aimed for human subjects, cannot be directly applied to them; rather, there is a need for delicate dose and therapeutic evaluations, as well as safety considerations. As a preliminary step to gather such pertinent information, animal studies are usually prescribed, though often with some limitations. There are several advantages of animal studies compared to human studies. Such studies can be conducted under relatively more controlled laboratory setups, and hence, statistical modelling and analysis aspects are relatively simpler. Also, cost-wise dosimetric studies are more economic than human ones. Moreover, from therapeutic and dose

setting points of view, such studies can provide valuable insights for subsequent studies involving human beings. On the other hand, subhuman primates may vary considerably from human beings in body metabolisms, pharmaceutical dynamics and kinetics, and hence, incorporating the findings from dosimetric studies for designing and modelling of toxicologic studies involving human beings may require a very careful appraisal.

Biological assays or bioassays

These are akin to dosimetric studies; traditionally, a new and a standard preparation (that could be toxic in some sense) are compared by means of reactions that follow their application to some subjects or living organisms, and the basic objective is to study the *relative potency* of the new preparation with respect to the standard one. In this sense, bioassays are somewhat different from toxicologic studies, where there could be a lack of a *control* or *placebo* group against which the *treatment* group should be statistically compared. Rather, in toxicologic studies, there is a more dominant dose–response relationship emphasis in which the dose levels may not be determined that accurately (nor the responses), and there could be more unaccountable variation due to the complexities of environmental and behavioural factors. Nevertheless, in terms of statistical methodology, bioassays in a broad sense[4] have some common features with toxicologic assays. Radioimmunoassays[5] are good examples in this context. Bioavailability and bioequivalence trials are also used for toxicologic studies[4].

Whither dose-response regression?

Although there is a predominant dose-response flavour in the sense that the dose refers to the various toxic substances and their prevailing levels while the response relates to the reactions that follow the intake of toxic material by various organisms, most notably, human beings, there are some crucial steps that need to be examined before attempting to use standard dose–response regression models. Let us iterate some of these roadblocks and explain the surrounding complexities and difficulties in the light of statistical perspectives.

(i) *Accurate or reliable identification of the actual toxins and their effective dose levels*: In general, there may be a (large) number of toxic elements (some of which may even be latent) with rather diverse mode of propagation, and hence, putting such a complex dose system may need considerable appraisal from various aspects. From a statistical standpoint, there may be the basic question of identifiability of all or even the majority of toxins that are particularly relevant in a specific context. In a majority of cases, even a simple model with a definite pattern of synergism of these toxins may not totally meet the light of reality of the actual pattern, and hence, elimination of redundant statistical information is a basic problem in such modelling and analysis schemes.

(ii) *Measurement errors and compliance issues*: A basic requirement in a dose-response regression analysis is the precise determination of the dose levels and the responses, no matter whether they are quantitative or quantal. However, in toxicologic studies, even in simple dosimetric ones, an accurate determination of the dose levels may not always be possible. In conventional linear models, measurement errors have been accounted in various forms[12], and it might be quite intuitive to incorporate some of these findings in the current context. However, as we shall see later on, there are some limitations for such standard adoptions. In dosimetric studies and more noticeably in actual environmental toxicology studies, an administered or recorded dose level may be grossly inappropriate as a regressor, and because of differential exposure levels, stochastic compliance error models[13] are more appropriate. Basically, such models are closely based on either Bayesian methodology or historical control and additional information on the compliance pattern.

Whither conventional regression models?

One of the basic approaches to the dose response regression modelling is to incorporate a suitable linear regression model, following transformation on the dose and/or response variables as needed to induce more linearity of the regression function and/or normality of the error components. In the same vein, generalized linear models[14] have been advocated for quantal response models and other types of nonstandard regression forms. In the context of multiresponse models with possibly non-continuous variables and latent effects, the appropriateness of (generalized or general) linear models may often be questionable. In fact, their primary shortcoming may be the lack of robustness property (to plausible departures from model-based assumptions, even to a moderate extent). On top of that, measurement or compliance error can introduce considerable complications in such standard models, and may even invalidate the models. For this reason, nonparametrics have been advocated for a wide class of models that arise in dosimetric studies, bioassays and bioequivalence trials which are potentially related to toxicologic studies[15]. For some allied semiparametric approaches, we may refer to Clegg *et al.*[16]. Model misspecification and measurement errors in covariates have been studied in this context (see also Clegg *et al.*[17]). Basically, these are all large sample procedures, and there remains considerable scope for developing moderate to finite sample solutions.

Spatio-temporal models

Generally, in environmental settings, data are collected on a set of grid-points scattered in a region, and a set of time

points covering a period. The matrix of observations is then incorporated in a spatio-temporal setup to draw conclusions for the region and time period. Since, we may have generally count variables, or even qualitative data models, without a reasonably stationary pattern, the usual *variogram*-model based statistical analysis may not be suitable; in fact, the very linearity or additivity of the accountable factors might be questionable, and hence, the usual methods based on *kriging* and spatial homogeneity[18] may not be that appropriate. Actually, some local homogeneity of spatial covariation is often assumed in order that standard kriging and variogram models can be adapted, though with a slower rate of convergence. Bayesian methodology, under appropriate priors on the spatial variation, is also being applied in this context[19]. Small area estimation procedures are also being incorporated (ref. 20, and references cited therein). A complete nonparametric approach may usually require larger sample sizes. In this respect, nonparametric and semiparametric longitudinal and spatial data analysis models (viz. ref. 3, and references cited therein) appear to be more appropriate. This is currently an active area of statistical research, and there is ample room for further developments.

Stochastic partial differential equations

In the context of PBPK models and in pharmacodynamics, some work has been in progress wherein SPDE have been incorporated with Gaussian white noise components. This is a promising line of attack, and much of the success would probably depend on how we can have sound statistical interpretations of the Fourier coefficients that appear in these SPDEs, and how effectively we can estimate these coefficients from experimentally acquired or observational data sets not enormously large in size.

1. Sen, P. K., *Encyclopedia of Environmetrics*, Wiley, UK, 2001a (in press).
2. Sen, P. K., *Encyclopedia of Environmetrics*, Wiley, UK, 2001b (in press).
3. Sen, P. K., in *Handbook of Statistics*: *Bioenvironmental and Public Health Statistics* (eds Sen, P. K. and Rao, C. R.), Elsevier, Amsterdam, 2000, pp. 247–324.
4. Sen, P. K., *Encyclopedia of Biostatistics* (eds Armitage, P. and Colton, T.), Wiley, London, 1998, vol. 1, pp. 372–384.
5. Finney, D. J., *Biometrics*, 1976, **32**, 721–730.
6. Hulka, B. S. and Margolin, B. H., *Am. J. Epid.*, 1992, **135**, 200–207.
7. Sen, P. K. and Margolin, B. H., *Sankhyā, Ser. B*, 1995, **57**, 252–276.
8. Durbin, R., Eddy, S., Krogh, A. and Mitchison, G., *Biological Sequence Analysis*, Cambridge Univ. Press, London, 1998.
9. Fisher, R. A., *Design and Analysis of Experiments*, Oliver and Boyd, Edinburgh, 1935.
10. Roy, S. N., *Some Aspects of Multivariate Statistical Analysis*, Statistical Publ. Soc., Indian Statistical Institute, Calcutta, 1957.
11. Sen, P. K., *J. Stat. Plann. Infer.*, 1999, **75**, 433–451.
12. Fuller, W., *Measurement Error Models*, Wiley, New York, 1987.
13. Chen-Mok, M. and Sen, P. K., *Commun. Stat. Theory Methods*, 1999, **28**, 931–946.
14. McCullaugh, P. and Nelder, J., *Generalized Linear Models*, Chapman-Hall, London, 1989, 2nd edn.
15. Kim, H. and Sen, P. K., *Sankhyā*, 2000, **B62**, 119–132.
16. Clegg, L. X., Cai, J., Sen, P. K. and Kupper, L. L., *Sankhyā*, *Ser. B*, 2000, **62**, 25–42.
17. Clegg, L. X., Cai, J. and Sen, P. K., *Biometrics*, 1999, **55**, 805–812.
18. Piegorsch, W. W., *Handbook of Statistics*, *Environmental Statistics* (eds Patil, G. P. and Rao, C. R.), Elsevier, Amsterdam, 1994, vol. 12, pp. 535–559.
19. Dey, D., Muller, P. and Sinha, D. (eds), *Practical Nonparametric and Semiparametric Bayesian Statistics*, Springer Lecture Notes in Statistics, No. 133. Springer, Heidelberg, 1998.
20. Datta, G. S., Ghosh, M. and Waller, V., in *Handbook of Statistics: Bioenvironmental and Public Health Statistics* (eds Sen, P. K. and Rao, C. R.), Elsevier, Amsterdam, 2000, vol. 18, pp. 223–245.
21. Dungworth, D., Kimmerle, G., Lewkowski, J., McClellan, R. and Stober, W. (eds), *Inhalation Toxicology: The Design and Interpretation of Inhalation Studies and their use in Risk Assessment*, Springer-Verlag, New York, 1988.
22. *Guidelines for Reproductive Toxicity Risk Assessment*, US EPA Office of Research and Development, Washington, DC, 1994.
23. Sen, P. K. and Rao, C. R. (eds) *Handbook of Statistics*, *Bioenvironmental and Public Health Statistics*, Elsevier, Amsterdam, 2000, vol. 16.