International Journal of Pharmaceutics, 8 (1981) 73–80

Review Article

PROTOCOLS FOR STABILITY TESTING

S.A. HILL and K.A. KHAN

Beecham Pharmaceuticals, Research Division, Clarendon Road, Worthing, West Sussex BN14 8QH (England)

(Received February 9th, 1981)
(Accepted February 19th, 1981)

SUMMARY

Protocol planning for stability testing is influenced by several factors such as physicochemical properties of the drug, type of product, regulatory requirements of the intended market and the proposed uses of the data obtained. In addition to specific job requirements, the priority allocation and the availability of resources also affect the protocol design. This review describes the objects and design of protocols for stability evaluation with particular emphasis on relatively unstable compounds. The relevance of stability testing to various stages of product development is examined. Problems of shelf-life assignment, preparation of samples for inspection, use of controls and extrapolation of results are also discussed.

INTRODUCTION

Consumer acceptability of pharmaceutical products relies to a considerable extent on the retention of their necessary qualities throughout a reasonable shelf-life. Retention of quality over shelf-life must be demonstrated, hence the need for stability tests. Every product is subjected to a series of such tests during its evolution, and the objectives, and hence the protocols, at each stage in development and marketing differ in important respects. However, the prime objects of the test programme must be the assurance of product quality to the patient in respect of efficacy and safety (general terms which encompass such qualities as potency, purity and bioavailability). An inherent danger in devising general protocols is that they may be applied indiscriminately without regard to factors specific to individual products. This can lead to resource wastage, and needless delays in development programmes. All general protocols are guidelines only, and the necessity for every item included needs to be carefully assessed when individual protocols are drawn up. Factors influencing the selection of items for inclusion comprise physicochemical properties of the drug, the type of product, regulatory requirements for intended markets, and proposed uses for the data obtained. The tendency is to make

the test as extensive as possible, and skill and realism are essential in matching the design to the allocated priority and hence to the resources available. Such skill is necessary, for example, in attempting to foresee problem areas and weighting the tests accordingly.

General considerations relevant to protocol design include the availability of analytical techniques and batch sampling methods. The identification of routes of decomposition on ageing and development of methods for quantitatively analyzing samples for active ingredients and degradation products must be initiated as early as possible in the programme. Chromatographic methods are currently widely used, especially HPLC, and the presence of degradation products can often be readily detected by such separation techniques. However, characterization can be difficult and tedious, in particular the preparation of reference compounds sufficiently pure for quantification. In the early stages of development, if the precision of available methods is low, adequate replication is vital. Batch sampling is important in validating batches, that is in assuring consistency within and between test samples.

STABILITY TESTING DURING VARIOUS STAGES OF PRODUCT DEVELOPMENT

The phases in product evolution may be defined as follows: (1) research; (2) preformulation; (3) full development; and (4) post-marketing.

(1) Research

Little stability testing is undertaken in the basic research phase, although some work may be carried out to assign limited shelf-lives to research compounds to assist the biologist and pharmacologist. Preliminary stability information may also be obtained during the investigation of assay methods.

(2) Preformulation

The programme of formal stability testing begins in this phase and expands in a manner and at a rate dependent on factors such as ease of compound preparation, and overall priority of the project. The main objectives of stability tests in the preformulation phase are the following.

(a) Selection of the preferred form of the compound. A decision on which form to develop (e.g. which salt, or which physical form) is required at the earliest possible time, to avoid resource and time wastage with unsuitable candidates. Reaching a decision may be quite complex, since the many factors relating to preparation and subsequently manufacture of the compound must be considered against those concerned with its formulation, ease of dosage form manufacture and stability. Characterization of candidates will include a heat-accelerated stability test. The test is likely to be constrained by limited supplies of material and assay methods of doubtful specificity. However, the objective is to make a comparison between several possibilities, and comparative rather than absolute results are all that are required. Batch homogeneity is important, and the materials should be well-blended before subdivision into test samples. The possible use of simple physical tests for evaluation should not be discounted. Colour changes, for

example, may be a good indicator of comparative stability.

(b) Screening for significant environmental challenges. These challenges, which include heat, light, humidity, microbiological contamination, etc., can all be enhanced to accelerate decomposition. Instability under enhanced stress does not automatically imply unacceptable performance under real life conditions, but the converse is nearly always true. Care is needed in enhancing thermal stress, since the adoption of a wide temperature range, or temperatures much above ambient, can lead to pitfalls such as changes in decomposition pathways and kinetics. Care is also necessary with light challenge, which may induce a concomitant rise in temperature.

Some typical protocols for tests challenging a compound in the dry state with temperature and humidity might be: temperature: 3 months at 50°C, 1 week at 80°C; humidity: 3 months at 30°C/75% RH (in open dishes). For compounds intended for use in solution (e.g. injectables), pH may be important and a range can be tested at appropriate temperatures. Similar tests can be applied to suspensions. Other challenges can be incorporated. as suggested, for example, by Akers (1976).

(c) Checking the stability of preparations for toxicological and pharmacological studies. Whilst these preparations are usually formulated very simply to facilitate administration to the test species, their stability has to be demonstrated. This is now a requirement in GLP. For compounds administered by injection, stability to autoclaving must be checked. Long-term storage tests prior to use may be impractical, and in these circumstances short shelf-lives are arbitrarily assigned, usually 3–6 months, and the materials are stored under refrigerated conditions. There may be no detailed protocols for stability studies on such supplies, and it is usual to check and revalidate them on their reaching the originally assigned expiry date.

The above objectives are inherent in the preformulation phase. Others are ideally incorporated at this stage, but may be limited in scope owing to a low priority or resource constraints. These objectives are the following.

(d) Investigation of the kinetics of potency loss and the feasibility of prediction from accelerated data. Knowledge of the kinetics of compound degradation can be of considerable value in product development, particularly for stability prediction and elucidation of degradation pathways. A detailed kinetic study of the active compound should be undertaken as soon as sufficient and appropriate analytical methods and material supplies are available. Depending upon project priority, this may occur in the preformulation phase, or after the compound has entered full development.

Accelerated stability testing

The accelerated temperature studies undertaken to assist selection of compound form will provide data valuable in setting up studies optimally designed for kinetic analysis. For these experiments, assay specificity, and high accuracy and precision are mandatory. The degree of precision will dictate the level of assay replication, which will probably be greater for the initial determination than subsequently since this is the reference point for all later assays. To apply the Arrhenius coefficient procedure, at least 3 and preferably 4 or 5 temperatures must be employed, and these should be approximately evenly spaced. The highest temperature sample should lose 70–80% of its initial potency in a week. The validity of a prediction can of course only be tested in retrospect, and to

enable this to be done, the low end of the temperature range will be that covered by the standard long-term storage test schedule (e.g. 5°C, 20°C, 30°C and 37°C). Other temperatures will be fitted in between these limits as appropriate. The routine inspection schedule requires a greater inspection frequency early in the storage period, since it is common for potency loss to occur most rapidly in this phase. Sample sizes are governed by the number and quantity requirements of the tests to be applied. The mathematical treatment of stability data used for prediction is well documented (Ferguson, 1978). Statistical aspects are important, and have been extensively discussed in the literature (Bentley, 1970; Clark and Hudson, 1968; Davies and Budgett, 1980). Examples of problem areas include the uncertainty in curve-fitting, variations in assay error, and errors resulting from data transformation. Compounded error factors such as these can result in wide confidence limits to predicted shelf-lives.

(e) Testing the compatibility of the compound with excipients that may be used in its formulations. These studies are devised to detect excipients which will or may result in poorer stability of formulated products than of the active material. They are usually heat-accelerated and complement differential thermal methods. A factorial design for such studies has been suggested by Leuenberger and Becher (1975).

CTC and IND submissions

Data obtained in the preformulation stage will be used to support IND and clinical trial registrations, and the relevant regulatory requirements must be incorporated in the protocols (Mollica et al., 1978) (also see the U.K. D.H.S.S. 'Notes for Guidance on Applications for Product Licences and Clinical Trial Certificates', 1977).

Some companies adopt a policy of developing in the preformulation stage what are termed 'service forms' (Worthington, 1979). These are dosage forms produced in a limited development programme to ensure early delivery of material for human trials. They are not fully optimized formulations; only very limited stability data on them may be available at the time of clinical trial registration and a commitment to provide stability results to the regulatory bodies is made.

(3) Full development

When the compound enters full development, raw material scale-up proceeds, frequently accompanied by process changes. Each of these has to be reviewed, and its possible significance to stability assessed. If considered necessary, raw material prepared by the new process must be storage tested along the lines already discussed, with material made by the standard process as control. The problem of within-batch variation may assume greater importance on scale-up, and each batch should be milled or screened and well-blended before sampling and subdividing.

Product formulation proceeds in parallel with compound process development and replicate formulations are often devised as an insurance. Three batches of each candidate formulation are tested, in the pack or packs intended for marketing. Formulation stability tests usually include raw material controls, and these provide useful data to supplement earlier tests. Further batches of the formulation are tested when the product is scaled up. The importance of processing in final product quality varies a great deal

from one formulation to another. At one end of the scale, injectables are prepared by readily standardized processes. At the other, many tablets, and especially liquid products such as emulsions and creams, are highly dependent on process variation and stability is one aspect of product quality that may be affected. For these products, limited stability testing is an economically justifiable part of process optimization and validation. Test conditions must take the type of package into account; any packs for which the moisture protection properties are not tried and proven must be tested against humidity stress. Temperature/humidity combinations that may be used include 20°C/75% RH and 30°C/75% RH. Cycling or alternating conditions can also be applied.

The inspection and evaluation of formulations involve many more tests than the assays and visual inspection applied to raw materials. For compliance with GLP, all analytical methods must be specified and fully-documented, or referenced to an official compendium when appropriate.

Special tests that may be required for inclusion in the protocol include 'use' and 'travel' tests, and product/pack interaction studies.

The data generated in this phase of the product life constitute the bulk of that submitted in support of product registration for marketing. The protocol must therefore be checked to ensure that the data collected will be appropriate and sufficient to meet the requirements in intended market countries.

Shelf-life assignment. Calculation of and support for over-ages and shelf-lives are major uses of formulated product stability data. 'Shelf-life' may be defined as 'that period after manufacture during which a batch of the product will remain within the specified potency range, *at an acceptable level of probability'*. The 'acceptable level of probability' may vary from one market to another, and both over-ages and shelf-lives may therefore also vary, even when derived from the same set of stability data. The main difficulties experienced in allocating over-ages and shelf-lives in the pre-marketing phase are firstly, the small number of batches for which stability data are available; secondly, the uncertain relationship between controlled temperatures selected for storage testing and variable conditions existing in the relevant markets (the concept of 'virtual temperature' (Haynes, 1971) is useful, but has practical drawbacks (Scher, 1980); and thirdly, the uncertainty in extrapolating from short-term data. Assignment calculations are also influenced by specification data, official monograph requirements where relevant, and the manufacturing tolerance for the product.

(4) Post-marketing

Stability tests in the post-launch phase comprise further investigation of compound and formulation process changes, formula modifications, etc. Additionally, production batches of products are stability tested.

Some aspects of product stability may not be finalized by product launch, and need to be clarified subsequently. With moisture-sensitive products, for example, setting a moisture specification may require extensive tests to accurately define the critical content. Formula modifications may be necessary to improve unsatisfactory features of the product, or they may be imposed by alterations in raw material specifications or availability. Process optimization and validation is another aspect of formulation development that frequently continues post-launch.

An important objective of post-marketing storage tests is confirmation of over-ages and shelf-lives, and support for extension of the latter. Only when data on sufficient batches ove a sufficient time span are available, can meaningful estimates of the probabilities of batches reaching a particular shelf-life with a particular over-age be made. These calculations permit a range of options for marketing departments.

SOME STABILITY TESTING PROBLEMS AND THEIR EFFECT ON PROTOCOL DESIGN

(1) Selection of temperature to represent conditions in the market place. This has already been mentioned as a difficulty in the assignment of shelf-lives. Test conditions need to be tightly controlled to minimize errors in prediction calculations. In real life, variable conditions are the norm, with the possible exception of air-conditioned warehousing. There is no substitute for stability testing in the actual field conditions the marketed product will experience, even though the difficulties in selecting 'typical' conditions in any one market are again obvious and legion. In laboratory tests, experience has shown that very crude correlations can be adequate for initial shelf-life assignments, and relating continuous 20–25°C to temperate markets, and 30°C to tropical is quite widely accepted by both manufacturers and registration authorities. For moisture-sensitive products in susceptible packs, humidity-controlled storage is employed, and 20°C/75% RH is appropriate for temperate conditions and 30°C/75% RH for most tropical environments.

(2) Preparation of samples for inspection. An area requiring special attention is the preparation of samples for inspection, in order to ensure uniformity and to permit valid comparison of the results of different tests on the same sample (for example, potency and water content). The necessity of blending bulk substances before subdividing into test samples has already been mentioned, and reblending each individual sample at the time of inspection is equally essential, unless the entire sample is used in the assay. The same principle should be applied to powder blend formulations (such as powders for reconstitution into oral or parenteral solutions or suspensions). In most instances, it is better to assay after reconstitution, since this avoids problems such as powder homogeneity or demixing and preferential adhesion of the active substances to glass. Adsorption of drug or excipient to plastic packaging may also present problems. Solid dosage forms often require more than one unit for assay and should be reduced to powder form by grinding or emptying from the shell as appropriate and blending before sampling for assay.

(3) Controls. Raw material in sealed glass tubes may serve as a control for solid dosage forms. Blank preparations, identical to the actual product except for omission of the active ingredients, are often useful as controls especially for creams, suspensions and lotions.

Since a majority of compounds are stable at 5°C or below, samples stored at these temperatures are often used as controls representative of the initial condition. Increases in the degradation products can often be estimated by comparing high sensitivity HPLC profiles of the test sample with that of the 5°C sample. A comparison of the main component result on a 5°C sample with the initial value is also a useful check on assay validity and reproducibility.

(4) Physical stability. Our knowledge of the mechanisms responsible for physical deterioration of products on storage, and methods used to evaluate such changes, are rudimentary in comparison with those relating to chemical stability of the active ingredient. In many cases the accelerated tests used to predict chemical degradation are not applicable for physical stability testing with the exception perhaps of exposure of solid dosage forms to high humidity. Even though the active ingredients of a product may not be subjected to degradation by excess humidity, the physical integrity of tablets or capsules may be impaired, causing changes to parameters such as appearance and dissolution. In some instances, chemical degradation is accompanied by physical changes or may be attributed to pack deficiencies and this is a good reason for performing the physical inspection prior to assay. The subject of physical stability testing has recently been reviewed by Rhodes (1979).

(5) Microbiological stability. For a multi-dose liquid preparation effectiveness of the preservative should be evaluated as part of the stability testing programme. The stability protocol should be designed to confirm by both assay and by microbial challenge that an effective level of the preservative is maintained (Moore, 1978).

(6) Extrapolation of results. Products are frequently marketed at a range of strengths, and in a range of pack sizes. Is it necessary to carry out a full storage test on every one? Some general comments are possible; for example, tablets or capsules of different strength made from the same basic powder or granule mix will rarely exhibit significant stability differences, provided the equipment used to manufacture them is also similar. (Nevertheless, some countries do require results at all strengths for registration.) As regards containers, extrapolation from one exhibiting good protection from the atmosphere to one which is less good should be acceptable. However, this again needs checking for registrability. Also in this connection, extrapolation between different sizes of containers has been considered by the FDA, and a guideline given (Banker, 1978).

Stability testing is probably the greatest consumer of resources of all the aspects of product formulation, and it is vital that the tests be as cost-effective as possible. Protocols should be drawn up to meet clearly defined objectives, and where they are based on general schemes, inclusion of each item should be carefully considered and adequately justified.

REFERENCES

Akers, M.J., Preformulation testing of solid oral dosage form drugs. Can. J. Pharm. Sci., 11 (1976) 1–10.

Banker, G., Comments on validation and Expiration Dating, International Conference on Pharmaceutical GMP in the Pharmaceutical and Related Industries, London, 5–7 December 1978.

Bentley, D.L., Statistical techniques in predicting thermal stability. J. Pharm. Sci., 59 (1970) 464–468.

Clark, C.J. and Hudson, H.E., Accelerated stability tests, practical considerations in their conduct and analysis. Manuf. Chem. Aerosol News, 39 (1968) 25–27.

Davies, D.L. and Budgett, D.A., Accelerated storage on pharmaceutical products, effect of error structure of assay and errors in recorded temperature. J. Pharm. Pharmacol., 32 (1980) 155–159.

Ferguson, L.D., Evaluation of drug stability, FDA By-Lines, 8 (1978) 281–320.

Haynes, J.D., World wide virtual temperature for product stability testing. J. Pharm. Sci., 60 (1971) 927–929.

Leuenberger, H. and Becher, W., A factorial design for compatibility studies in preformulation work. Pharm. Acta. Helv., 50 (1975) 88–91.

Mollica, J.A., Ahuja, S. and Cohen, J., Stability of pharmaceuticals. J. Pharm. Sci., 67 (1978) 443–465.

Moore, K.E., Evaluating preservative efficacy by challenge testing during the development stage of pharmaceutical products. J. Appl. Bacteriol., 44 (1978) 43–55.

Notes for Guidance on Application for Product Licence (MAL 2) and Clinical Trial Certificate (MAL 4), Medicines for Human Use, issued by Department of Health and Social Security, U.K., 1977.

Rhodes, C.T., Physical stability of pharmaceutical products. Drug. Dev. Ind. Pharm., 5 (1979) 573–587.

Scher, M., Kinetic rates as a parameter for product stability calculations. J. Pharm. Sci., 69 (1980) 325–327.

Worthington, H.E.C., Some problems in the U.K. with the provision of new drug dosage forms for pre-marketing tests in man. Drug Dev. Ind. Pharm., 5 (1979) 245–262.