Comment on sodium cromoglicate: an ineffective drug or meta-analysis misused? by Stevens et al., Pharmaceutical Statistics

John A. Lewis*†
School of Mathematics and Statistics, University of Newcastle upon Tyne, 1 Park Terrace, NE1 7RU, UK

Stevens et al. [1] have reviewed a series of meta-analyses concerning the use of sodium cromoglicate (SCG) in childhood asthma, each carried out by a largely identical group of authors. Stevens et al. believe that the interpretations of these meta-analyses are all too strongly negative and that this has subsequently been reflected in international guidelines and consensus statements. They say that, as a result, the extent of use of a valuable drug in children may now be inappropriately low. There are three important strands in this debate.

OLD DRUGS

First, the drug in question is relatively old, having entered clinical trials in adults about 40 years ago. Unusually (and commendably), trials in children started very early in development. Hence any review of the totality of evidence has to cover trials carried out against a background of widely differing methodological standards and medical practices. The later trials would also have had a range of new objectives concerned, for example, with the efficacy of new formulations and methods of delivery and with efficacy in clinically important subgroups, rather than re-establishing the basic efficacy of the drug. This scenario poses special difficulties for meta-analysts.

Despite the general improvement in clinical trial methodology since the 1960s, many older trials, particularly those of treatments for symptomatic diseases, were adequate for distinguishing active treatments from placebo in a manner that could be supported by observations made in subsequent clinical use. SCG was such a treatment. The mass of evidence concerning this drug, using a variety of outcomes in numerous trials and in practice, has left little doubt that it is, in broad terms, an effective and safe agent. That is why it was licensed for the treatment of asthma throughout the world.

Any meta-analysis of a quantitative outcome gathered from trials carried out over such a long period of drug development in a disease such as asthma is almost bound to encounter heterogeneity. Stevens et al. describe how designs of the controlled trials of SCG were modified for changing objectives over the years. There are explanations of the reasons for studies with a common background treatment (isoprenaline), studies in different age groups of children, and studies of alternative formulations and delivery systems for children and for very young children. Many of the original individual trials, when examined separately, would have been expected to be capable of providing evidence of the differing size of clinical benefits when the drug was used in these different ways and different populations. Differences in efficacy were indeed found. This background would be vital for the successful design and interpretation of a meta-analysis, both as a guide to the trials that might be profitably combined, and as a guide to the interpretation of any conclusions. Indeed a full examination of the background might deter some meta-analysts from proceeding.

META-ANALYSIS METHODOLOGY

The second strand in the debate is the methodology of the meta-analysis itself. Questions are raised concerning the choice of primary outcome(s) and the relative weight to be given to evidence derived from other outcomes. The identification and selection of studies has not been straightforward. Predictably, heterogeneity has been prominent in the analysis.

Meta-analysis has come a long way since the heady days of the 1970s and 1980s. Although these decades were by no means
the start of meta-analysis, the successful uses at this time in cardiovascular disease and cancer gave it a considerable impetus. Early successes in these disease areas were largely concerned with mortality outcomes or serious morbidity outcomes and resolved long-standing debates about the value of some medical interventions such as the use of beta-blockers post infarction [2] and the use of tamoxifen as adjuvant therapy in breast cancer [3]. Meta-analyses at that time tended to be large co-operative enterprises conducted by collaborative groups of clinicians and statisticians representing much of the key research work in the relevant area of medicine. The methodology used was relatively simple.

One of the major and heart-warming events associated in time with this impetus was the creation of Cochrane Collaboration (CC). This ended the era of the opinion-led review of medical interventions and brought the quantitative review to the centre of the action. Huge positive strides were made as evidence-based medicine was born. It is vital to remember what the world of overviews was like before quantitative methods were properly and regularly used.

The reported meta-analyses of SCG are considerably less straightforward than the early cardiovascular ones. For example, there is no obviously dominant outcome to parallel mortality and the information provided by a number of outcomes would be expected to carry some weight in drawing conclusions. Under these circumstances it seems inappropriate to put major emphasis on the meta-analysis of a primary outcome that is reported in very few of the trials, as was done in one of the reported meta-analyses [4]. This amounts to saying that most of the earlier studies are not able to provide reliable evidence for or against efficacy. This may be true but, if so, it should be argued directly rather than implied by the limitations on a meta-analysis.

The methodology employed in the meta-analyses of SCG has been chosen specifically to deal with the encountered heterogeneity. Qualitative outcomes often provide adequate power for the detection of heterogeneity in a meta-analysis and there should be no surprise about the heterogeneity of efficacy in these meta-analyses, even within the restricted population of children. The interesting question is how to interpret it. In some of the reported meta-analyses, the use of tolerance intervals is described, a technique whose purpose does not seem to fit the situation and whose interpretation in this context is obscure. Random effects models are also extensively used. When the basis of the heterogeneity of the treatment effect is known or strongly suspected, it is much better to explore it and describe it rather than incorporate it into a random element of the model. Use of random effect models may mislead the analyst into claiming that no treatment effect has been detected. If the reason for using the random effects model is that statistically significant heterogeneity has been detected, then the drug under examination is clearly not a placebo and most certainly does have an effect. But it has an effect size that varies according to the circumstances. So once heterogeneity has been identified, therefore, interest should focus on the circumstances under which this variable treatment effect is largest. That way it is possible to find results that will beneficially influence the future use of the drug. Retreating into a random effects model alone leads to a loss of useful information.

A further point can be made about the random effects model in the CC meta-analysis. Results have not been reported in a consistent manner across all variables but have been reported differently for each variable according to the results of the heterogeneity test for that variable. This defies common sense. If heterogeneity exists, then it will be there to a greater or lesser extent for all these associated measures of efficacy. So if it is to be accepted as a fact, then it should be accepted for all variables and not selectively.

Finally, the interpretation of heterogeneity should take into account what is known about the overall efficacy of SCG. It is not plausible that a drug that is effective in adults would be simply ineffective in children of all ages. There may be reduced efficacy and this reduction may be greater in younger children. Most of all, the conclusions drawn from a meta-analysis should be couched in terms that reflect their degree of certainty or lack of it. The meta-analyses of SCG seem to have drawn remarkably firm conclusions given the methodological difficulties that were encountered.

INFLUENCE ON USE

The third strand is the way in which these meta-analyses have affected public health practices. Has this happened in an appropriate manner and with sufficient checks and balances?

It is important to say that this strand of the debate should not be specifically about the CC. Only one of four publicly reported meta-analyses in question here was carried out under the auspices of the CC. So if (and I repeat ‘if’) there was any failure of vetting procedures for meta-analyses of SCG, then the lessons drawn should also apply to the journals and scientific conferences that were involved. But let us consider a broader context than this. Meta-analyses of medical interventions are carried out by a wide variety of medical scientists, largely to influence perceptions of the effect of that intervention. This may be through a publication but it may also be through inclusion in a regulatory submission for a marketing authorization. Regulatory experience in Europe [5] suggests that meta-analyses carried out by the pharmaceutical industry are by no means uniformly well motivated, conducted and interpreted. European guidance has been written in an attempt to improve their general standard [6].

However, in the case of regulatory applications there are formally appointed independent bodies (regulatory authorities) that are charged with protecting public health from misleading scientific work. In the case of published meta-analyses it is the task of the scientific community, professional bodies and individuals to ensure that contentious work does not become embodied in guidance or practice. Taking into account all the financial pressures and desire for personal prestige that process has been remarkably successful over the years. However, it
remains true that meta-analyses that are published in prestigious journals and those that appear under the auspices of the CC are likely to be accepted as authoritative by the majority of readers, and it is hard to counteract this strong influence on the occasions that it appears to be inappropriate. It is important to recognize that science does not proceed by publishing only indisputable truth. If that was the process, the journals would all be empty. Journals and others must carry on contributing to a free debate. So the onus lies with those who produce guidance on treatment, and those who make decisions about treatment, to judge correctly the value of the publications and other work that they read. The difficulty of this task should not be underestimated. Perhaps the CC would consider whether there is any further way in which they can help.

REFERENCES


