REVIEW ARTICLE

Review articles, systematic reviews and meta-analyses: which can be trusted?

Przemysław Ryś[1], Magdalena Władysiuk[1], Iwona Skrzekowska-Baran[2], Maciej T. Małecki[3]

1 HTA Consulting, Kraków, Poland
2 NovoNordisk, Poland
3 Department of Metabolic Diseases, Jagiellonian University, Medical College, Kraków, Poland

KEY WORDS

diabetes mellitus, meta-analysis, systematic review

ABSTRACT

A large number of scientific articles published every year requires from practicing physicians the ability to choose among them and to use secondary studies, such as guidelines, review articles, meta-analyses and systematic reviews. The aim of this article was to discuss basic differences between meta-analyses and systematic reviews. Meta-analysis is a mathematical method of pooling the results of several or more studies; a meta-analysis may be based on a systematic review, but this is not always the case. A systematic review is a multistage process aimed at the identification of all reliable evidence regarding a specific clinical problem. Systematic reviews make it possible to objectively address particular issues according to the current state of clinical knowledge and therefore constitute a reliable basis for clinical decision-making. An appropriate systematic review should include: 1) a defined clinical question, 2) pre-specified inclusion and exclusion criteria, 3) complex search for medical evidence sources according to a search strategy, 4) critical evaluation of reliability of identified clinical trials, 5) qualitative or quantitative data synthesis and 6) evidence based conclusions. These simple criteria, formulated by Cook et al. more than 10 years ago, allow to differentiate between a reliable systematic review and a "quasi-systematic" one, as well as between a reliable meta-analysis based on a systematic review and a potentially misleading meta-analysis without a systematic review.

Correspondence to:
Assoc. Prof. Maciej Małecki, MD, PhD, Katedra i Klinika Chorób Metabolicznych, Uniwersytet Jagielloński, Collegium Medicum, ul. Kopernika 15, 31-501 Kraków, Poland, phone: +48-12-421-37-94, fax: +48-12-421-97-86, e-mail: mmalecki@cm-uj.krakow.pl
Received: November 3, 2008.
Accepted: February 9, 2009.
Conflict of interest: none declared.
Pol Arch Med Wewn. 2009; 119 (3): 148-156
Translated by Marek Koprowski, MD, PhD
Copyright by Medycyna Praktyczna, Kraków 2009

INTRODUCTION Practicing physicians encounter thousands of clinical trials concerning their specialty published each year. At present there are >18 million publications indexed in the MEDLINE database alone.[1] It is estimated that ca. 25,000 randomized clinical trials are published annually[2] and their total number may even reach 1,000,000[3]. Such a large number of clinical publications requires skillful use of available sources of medical evidence by both scientists and physicians in order to choose the most valuable papers, providing complete and true answers to formulated questions. There is little doubt that this group includes original articles concerning primary studies. In case of therapeutic interventions they are mainly randomized clinical trials, although well-designed observational studies may also be a valuable contribution to science and clinical practice.[4,5]

However, there are often several or even more than ten clinical trials dealing with the same problem. These publications are sometimes heterogeneous and drawing any conclusions based on such studies must be preceded by a substantial analytical work. Such work could hardly be expected from practicing physicians. Secondary studies, such as review articles, meta-analyses, systematic reviews and guidelines, are therefore needed. Meta-analyses and systematic reviews are more and more often published in indexed journals. Unfortunately, these 2 categories are sometimes confused.

Narrative reviews and systematic reviews A systematic review is a multistage process aimed to identify all available and reliable publications regarding a specific problem, to evaluate their reliability and to compile available data in a quantitative (meta-analysis) or qualitative manner.[6] Actual work is usually preceded by the development of a detailed protocol specifying main directions and assumptions of the research process.[7]

An important issue discussed in this article is the distinction between a systematic review and a narrative review. This could be done according to a set of simple criteria formulated by Cook et al.[8] However, most problems encountered by users of secondary studies are associated with meta-analyses. In our opinion, these terms are often confused.

Meta-analysis is actually a statistical method used for evaluation of pooled data from at least 2 (usually several or up to 20) trials in order to obtain more precise average results.[9] From this point of view meta-analysis is just a statistical test, like the t-Student test or the χ^2 test for a single study. Reliability of a meta-analysis depends not on its statistical use, but – in the first place – on the inclusion criteria for primary studies. Usually it is preceded by a systematic search in medical databases, which allows to identify all credible original studies fulfilling a set of pre-defined inclusion criteria. If this is the case, conclusions drawn from a meta-analysis may be treated equally to those drawn from a systematic review, as the results obtained may be considered reliable and based on complete and up-to-date information.[10] Moreover, a systematic review often includes quantitative data analysis – a meta-analysis. Nevertheless, sometimes a meta-analysis is based on several original papers selected by the authors in an arbitrary or unknown manner. Such a meta-analysis does not fulfill the criteria for a systematic review and its results do not reflect complete medical evidence, but a fragment of it, often arbitrarily selected. Results of such review articles and meta-analyses based on them are sometimes contrary to the results of systematic reviews dealing with the same problem. For example, in 2004 a meta-analysis by Hanefeld et al. was published[11], which indicated a favorable effect of acarbose on a reduction in the risk of cardiovascular complications in type 2 diabetes mellitus. However, that paper was not based on a complete systematic review and eventually included only 7 clinical trials, both published and unpublished, the meta-analysis of which indicated reduced risk of myocardial infarction (hazard ratio [HR]: 0.36, CI 95%: 0.16–0.80) and all cardiovascular events (HR: 0.65, CI 95%: 0.48–0.88).

A year later authors cooperating within the Cochrane Collaboration developed a systematic review on the efficacy and safety of acarbose in type 2 diabetes mellitus, which did not confirm previous optimistic reports and indicated that there are no reliable studies evaluating the effect of acarbose on the risk of cardiovascular disorders. The discrepancy between conclusions of both reviews arose from the fact that Hanefeld et al. used the results of several unpublished studies, but made no critical evaluation and did not present the results of particular studies but only pooled data. The authors of the Cochrane Collaboration review decided that those data should not have been used in the analysis because they were not published and their reliability was not assessed in the meta-analysis by Hanefeld et al.[12]

Even a more spectacular example of a meta-analysis not based (in our opinion) on a credible systematic review was the work by Nissen and Wolski, which – due to controversial results – immediately gained publicity. The study reported that use of rosiglitazone was associated with a statistically significantly increased risk of myocardial infarction (odds ratio [OR]: 1.43, 95% CI: 1.03–1.98) and increased risk of cardiovascular death (OR: 1.64, 95% CI: 0.98–2.74).[13]

Nearly at the same time a systematic review was published by the Cochrane Collaboration, in which a tendency to lack of benefits from rosiglitazone was found in certain studies; however, the differences were not statistically significant and meta-analysis was not performed due to substantial heterogeneity of the results observed in the studies.[14]

In the paper by Nissen and Wolski lack of a comprehensive approach to the search in information sources was one of the limitations most widely discussed in the literature.[13]

The studies included in the meta-analysis were characterized by significant heterogeneity with respect to duration of treatment, comparators and the study design. Despite those differences statistical methods assuming homogeneity of the studies were used for the analysis of pooled data, making it easier to achieve statistical significance.[15]

Systematic reviews To differentiate between a systematic review and a narrative review several key criteria (TABLE 1) formulated >10 years ago by Cook et al. need to be met.[8] These include:

1 a clearly formulated clinical question
2 pre-defined inclusion and exclusion criteria, according to which studies are included in the review
3 comprehensive search in medical evidence sources
4 critical evaluation of included clinical trials
5 quantitative data synthesis
6 reasoning based on scientific evidence.

These criteria do not include all important elements affecting quality of the review; nevertheless, they form a base for further, more detailed assessment. Evaluating a systematic review one should also consider whether it is up-to-date, was based on a protocol[7] and how the studies were selected and data extracted.

Formulation of a clinical question A clinical question should precisely define the objective and scope of the analysis. It cannot be too general, e.g. "Are long-acting human insulin analogs efficacious?". A properly formulated clinical question should reflect the PICO (population, intervention, comparator, outcomes) format.[16] The question should specify the type of participants, the type of the intervention (or exposure) and the most important outcomes. Usually this is enough to define the problem clearly. Sometimes information

TABLE 1 Differences between a systematic review and a narrative review (based on Cook et al.[8])

Feature	Narrative review	Systematic review
question	often broad in scope	often a focused clinical question
sources and search	not usually specified, potentially biased	comprehensive sources and explicit search strategy
selection	not usually specified, potentially biased	criterion-based selection, uniformly applied
appraisal	variable	rigorous critical appraisal
synthesis	often a qualitative summary	quantitative summary
inferences	sometimes evidence-based	usually evidence-based

concerning the comparators and clinical trial design is also taken into account. Of course, it is important to make the question as short and concise (leaving details to be specified in the inclusion and exclusion criteria) as possible, but in the first place the question must be unambiguous.[17] In the example mentioned above the clinical question should therefore be formulated in a more detailed and "answerable" way (if this is of interest to us), e.g. "Does use of long-acting human insulin analogs in patients with type 2 diabetes mellitus decrease the risk of nocturnal hypoglycemia as compared to isophane insulin?" In the same way the following question must be considered too general: "Are angiotensin converting enzyme inhibitors (ACEIs) efficacious?" We should rather make it more detailed in order to reflect a specific clinical problem. For example, the question may be presented as follows: "Do ACEIs used in patients with ischemic heart disease contribute to decreased risk of myocardial infarction or cardiac death?" We may also investigate, "whether use of ACEIs decreases mortality in patients with circulatory failure?"

Inclusion and exclusion criteria The criteria for inclusion and exclusion of primary studies into/out of the review should be defined according to a well-formulated clinical question. Their precise definition should precede the search in medical evidence sources and the selection of studies for analysis. Usually this is done during development of a protocol, at an initial stage of the work. For the Cochrane Collaboration review protocols are mandatory and published as electronic documents. Publication of the protocol makes development of the analysis more transparent and therefore increases the chance to obtain credible and objective results.[7]

The inclusion criteria should specify a disease or condition under consideration and respective diagnostic criteria. Additional limits concerning specific subpopulations (age, gender, concomitant diseases, past myocardial infarction, the metabolic syndrome) are possible; however, these should be based on biological premises or data from other studies. Unjustified restrictions in inclusion criteria should be avoided since the results of such a limited review may not reflect clinical practice. For example: in a review concerning efficacy and safety of metformin in patients with type 2 diabetes mellitus restricting the treatment to monotherapy may be justified, while defining the upper age limit or exclusion of studies, in which patients suffered from concomitant arterial hypertension, may result in limited applicability of the results. Another important stage is a definition of the investigated intervention and the selection of comparators. In case of therapeutic interventions a dose of the medication (and sometimes the route of administration) should be specified. Of course, the intervention may be a specific drug (e.g. metformin) or a treatment protocol (intensive insulin therapy, continuous subcutaneous insulin infusion [CSII] etc.). For applicability of a systematic review selection of the comparator is important. The investigated intervention should be compared with an alternative currently used in clinical practice and not just with a placebo or a rarely used alternate intervention. In case of long-acting human insulin analogs in type 2 diabetes mellitus isophane insulin will therefore be a natural comparator, while comparison with ultralente insulin (practically not used in Poland) will be less important.

In type 1 diabetes mellitus CSII may be considered an alternative for use of long-acting analogs, while in type 2 diabetes mellitus such a comparison would be unjustified. Endpoints constitute another element that should be specified in inclusion and exclusion criteria. In the Cochrane Collaboration reviews it is recommended to take into account all clinically significant endpoints, mainly those important in making a decision whether a specific intervention should be used. However, it is not necessary to include all endpoints reported in studies when their practical significance is low or negligible.[18] Apparently, studies dealing with a clinical question concerning one specific selected endpoint, such as the risk of hypoglycemia and cardiovascular events in patients with diabetes mellitus treated with glyburide as compared to other anti-diabetic medications[19] or the effect of rosiglitazone on the risk of development of a malignancy [20], are also useful. The last element defined in the inclusion/exclusion criteria is the design of studies included in the systematic review. If the objective is to evaluate efficacy and safety of a therapeutic intervention, then the most credible design (provided that it is properly applied) is that of a randomized controlled trial (RCT); in general, most systematic reviews are based on this type of primary studies. However, in certain cases it is necessary to resort to observational studies. Such a situation may occur if data from RCTs are insufficient, the objective is detailed assessment of safety, or an RCT would be objectionable on ethical grounds. Due to lack of appropriate randomized trials evaluation of the effect of combination therapy with metformin and sulphonylurea on the risk of cardiovascular complications could be performed by Rao et al.[21] only as a meta-analysis of data from observational studies. Assessment of the effect of

TABLE 2 Electronic medical databases to be searched for systematic reviews according to the Polish Guidelines of the Agency for Health Technology Assessment in Poland

main databases	Medline
	EMBASE
	BioMed Central
other medical databases	CINAHL Database
	PsycINFO
	European Public Assessment Report (EPAR)
	Health Canada
	Netherlands Pharmacovigilance Centre Lareb
	The Uppsala Monitoring Centre
	Thompson Micromedex
other sources	references of included studies
	periodic safety update reports
	conference abstracts
	registries of clinical trials

TABLE 3 Search strategy applied in the Pubmed database for identification of randomized controlled trials regarding insulin detemir in type 2 diabetes mellitus (September 30th 2008)

Nº	Scope of search	Key words	Results
#1	population	diabetes mellitus	250963
#2		diabetic	137536
#3		diabetes	315169
#4		diabet[a]	333246
#5		non-insulin-dependent diabetes mellitus	62095
#6		NIDDM	53403
#7		#1 OR #2 OR #3 OR #4 OR #5 OR #6	334187
#8	intervention	detemir	214
#9		insulin detemir	189
#10		levemir	190
#11		#9 OR #10 OR #11	215
#12	methods[a]	random	172652
#13		random[a]	574673
#14		random allocation	62603
#15		randomised-controlled-trial	309981
#16		randomized-controlled-trial	309981
#17		RCT	3203
#18		#12 OR #13 OR #14 OR #15 OR #16 OR #17	575404
#19	search results	#22 AND #14 AND #21	65

a simplified filter was used for identification of RCTs; complete, comprehensive and methodologically checked filters are available from the SIGN (Scottish Intercollegiate Guidelines Network) webpage (http://www.sign.ac.uk/methodology/filters.html)

cigarette smoking on development of type 2 diabetes mellitus is possible only by means of meta-analysis of observational studies, since – due to ethical reasons – no experimental trials could be performed.[22]

Comprehensive search in medical evidence sources The main objective of a systematic review is identification of all credible clinical studies dealing with a specific problem, defined by the clinical question. It is mandatory to perform the search at least in the most important electronic medical databases, including MEDLINE, EMBASE and CENTRAL. The former 2 are bases of abstracts published in the indexed journals. They include a vast majority of significant papers published worldwide, in English and in other languages. A characteristic feature of these bases is the fact that if a specific journal is indexed in a given database, then all articles published in it are listed in this base. However, sometimes (although rarely) good and credible studies are published in less recognized journals, indexed neither in MEDLINE, nor in EMBASE. This gap is filled in part by the CENTRAL database maintained by the Cochrane Collaboration – an independent, international, non-profit organization, whose mission is to provide reliable evidence in order to assist in decision-making – both in clinical practice and management of healthcare systems. Members of this organization cooperate within dozens of thematic groups, actively seeking for well-designed clinical trials in various sources of evidence. Identified trials are then registered in the CENTRAL database. It is important that search is not limited to generally available Internet databases (MEDLINE, EMBASE), but also includes journals not indexed in these bases, often published in original languages (other than English), and conference abstracts. The CENTRAL database contains therefore studies thoroughly selected from a wide range of sources. The bases described above constitute a "canon" to be searched during development of each systematic review; however, in certain situations it is necessary to search in specific medical databases or use other methods (reviews of references in the identified papers, contacts with the authors of the identified papers and experts in a specific field, and/or the manufacturer of the drug under consideration). Medical databases recommended by guidelines published in 2007 by the Agency for Health Technology Assessment in Poland are listed in TABLE 2.[23] One must remember that it is important not only to search through a specific medical evidence source, but to do it in the right way. For instance, when looking for information concerning efficacy and safety of insulin detemir in type 2 diabetes mellitus, it is not enough to write in the generic name. Usually it is necessary to construct more or less complex questions using a range of keywords; a set of such questions is called a search strategy. In general, its form reflects the clinical question, i.e. the PICO format. An example of a search strategy is presented in TABLE 3. In summary, searching in databases may be considered systematic, if the authors reported which medical databases were searched and what keywords were used.

Selection of clinical trials identified within a systematic review Searching through medical evidence sources in a manner described above results in identification of a large number of potentially

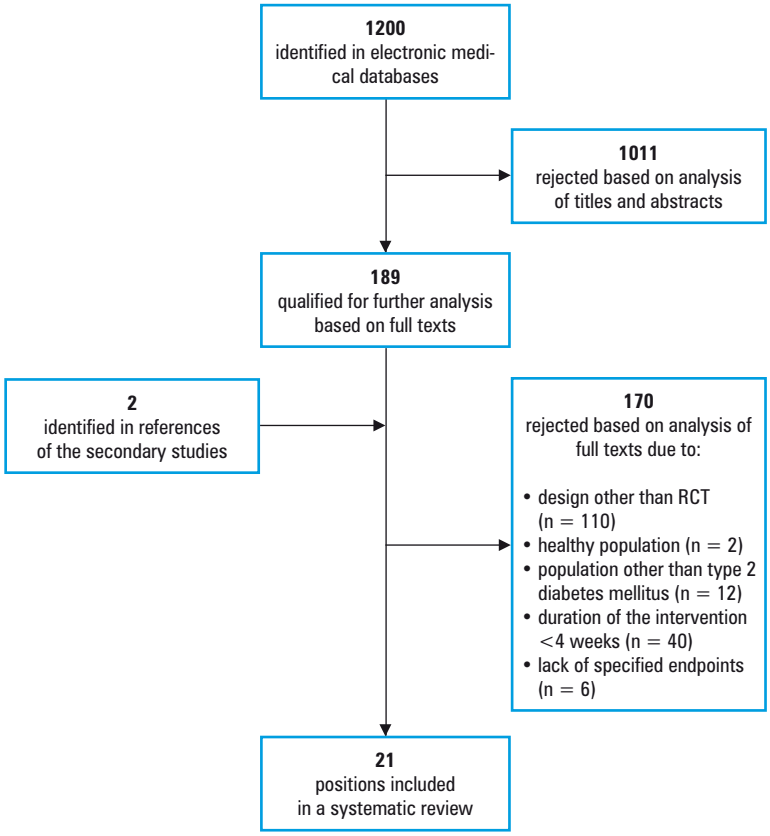

FIGURE 1 Selection of studies according to Quality of Reporting of Meta-analyses (a hypothetic example)

useful articles, of which a dozen or less (rarely more) will be included in the review. The following important stage is an appropriate selection of the identified publications according to pre-defined inclusion and exclusion criteria. This is done in 2 steps. First, initial selection based on titles and abstracts is performed: an article is rejected if its title or abstract is sufficient to consider the publication useless. The remaining papers undergo detailed verification based on their full texts. It is required that selection is done by at least 2 investigators working independently in order to decrease the risk of an error. In credible systematic reviews the whole selection process is presented using a special diagram according to Quality of Reporting of Meta-analyses (QUOROM), which makes it possible to repeat the process and verify its correctness.[24] An example of such a diagram is presented in FIGURE 1.

Critical evaluation of clinical trials identified within the systematic review Any conclusions concerning efficacy or safety of a specific intervention must be based on reliable scientific evidence. Not all studies identified within a review fulfill basic credibility criteria; the authors of a review are therefore required not only to identify relevant studies, but also to assess their quality. If studies are considered credible, conclusions based on their results (concerning efficacy or safety of a given intervention) may be drawn with confidence, while studies with methodological faults must be treated more carefully. The appraisal is usually based on structured scales or questionnaires. One of the simplest and most widely used scales for assessment of randomized trials is the scale proposed by Jadad et al. (TABLE 4)[25], that takes into account the presence and method of randomization, the presence of a double-blind design (and the method of blinding) and the completeness of observation (the number of patients lost from the study). This tool became very popular due to its simplicity and was widely used in the past; many authors still use this scale. It is also recommended in guidelines published by the Agency for Health Technology Assessment in Poland.[23] However, the Jadad scale was developed >30 years ago. These 3 decades have seen much development in methodology of randomized trials, their assessment and interpretation. The Jadad scale is therefore often inadequate for complete and appropriate evaluation of a study since it does not include a range of important elements affecting its credibility. The most recent Cochrane Collaboration guidelines concerning principles of development of systematic reviews recommend use of a somewhat more complicated tool, which makes it possible to take into account blinding with respect to the patients, the assessors and the analysts, intention-to-treat analysis and possible selectivity of the results reported.[26]

Meta-analysis – quantitative analysis of pooled data The results of identified studies must be presented in such a way as to make it possible to answer the formulated clinical question. The preferred form of presentation is quantitative analysis of pooled data, i.e. meta-analysis. However, in certain cases meta-analysis may not be feasible and descriptive methods (qualitative analysis) must be used.

Meta-analysis is a quantitative process in which data from single studies are pooled. Meta-analysis makes it possible to obtain one average value resulting from data from single trials concerning a selected endpoint and (as is the purpose of each statistic test) to answer the question whether observed differences may be due to pure chance. The weight of a given trial (i.e. its effect on the final result of meta-analysis) depends on the sample size and data scatter (precision of the results).

The final result is the difference between the intervention group and the control group expressed as a statistical parameter (e.g. relative risk, odds ratio, weighted mean difference etc.). Meta-analyses are typically presented in a graphic form called a forest plot or a meta-analysis plot. A vertical line in the central part of the plot, commonly known as the "zero effect" line, indicates lack of differences between the groups (FIGURE 2). Subsequent squares represent mean values reported in individual studies, while the ends of a respective segment indicate the confidence interval limits. The size of the square reflects the weight of the study (the larger the square, the greater effect of the study on final results of the meta-analysis). If a horizontal segment placed over a specific square crosses the "zero effect" line, then the difference between compared interventions is statistically

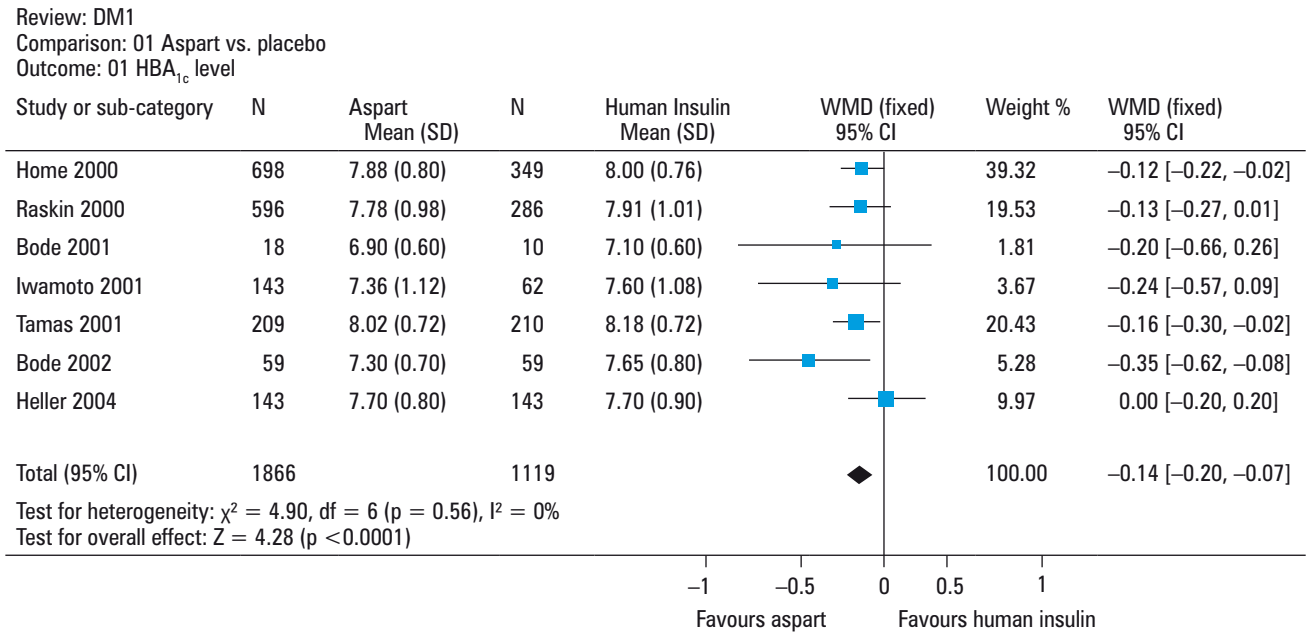

FIGURE 2 Weighted mean difference between insulin aspart and human insulin in type 1 diabetes mellitus
Abbreviations: DM1 – type 1 diabetes mellitus, SD – standard deviation, WMD – weighted mean difference (meta-analyses prepared in Review Manager v.4.2.)

insignificant (for a given p value – the threshold – and confidence interval). If, however, the whole segment lies on one side of the "zero effect" line, the difference is statistically significant.

The lowest part of the graph represents the result of meta-analysis (in the form of a diamond). The diagonal between the upper and lower apex of the diamond represents the average value, while the left and right apex determine confidence interval. Interpretation of clinical significance is the same as for a single study: if the diamond crosses the "zero effect" line, then the difference is statistically insignificant; if the whole diamond is located on either side of the line, then the difference is statistically significant.

An example of a result of meta-analysis concerning a comparison between insulin aspart and human insulin in type 1 diabetes mellitus with respect to weighted mean difference in glycated hemoglobin level is presented in FIGURE 2 (Pankiewicz, Ryś – unpublished data). Detailed analysis of this figure allows for the following observations:

1 seven studies were taken into account in the meta-analysis

2 in 3 studies statistically significant differences in favor of insulin aspart were demonstrated (respective segments are located to the left of the "zero effect" line)

3 the results of the remaining 4 studies are statistically insignificant (respective segments cross the "zero effect" line)

4 the result of the meta-analysis is statistically significant in favor of insulin aspart (the whole diamond is located to the left).

Clinical significance of the observed difference requires separate assessment. The result was presented as glycated hemoglobin (HbA_{1c}) level reduction, i.e. a biochemical parameter, not a measure of direct benefit to the patients. Demonstration of a relationship between this parameter and incidence of clinically important endpoints (e.g. late complications of diabetes) makes it possible to assess how much useful for the patient is the benefit observed. In the Diabetes Control and Complications Trial (DCCT) a relationship between HbA_{1c} reduction and development of microangiopathy was demonstrated; however, differences with respect to the HbA_{1c} observed in the DCCT study were higher than those presented in the figure – 1–2%.[27] It was not demonstrated whether a reduction of the glycated hemoglobin level by 0.1% was associated with reduced risk of complications in diabetes mellitus. On the other hand, the relationship between HbA_{1c} and long-term complications is linear and therefore it might be stipulated that each reduction in the HbA_{1c} level is beneficial for the patient.

Meta-analysis is therefore a way of data processing using statistical methods (analogically to the t-Student test, the χ^2 test or analysis of variance [ANOVA]) and reliability of a particular meta-analysis depends on quality and homogeneity of the studies included. The first and most important question to be considered is completeness and reliability of information used. As mentioned above, if a meta-analysis was preceded by a reliable systematic review and all credible information was included, then its results may be considered reliable and consistent with the current state of knowledge. Completeness of information is an important, but not the sole determinant of credibility of a meta-analysis. Prior to pooling the results the authors of a meta-analysis should consider whether differences between the studies with respect to population characteristics, the intervention, the comparator (a control group) or methodology of the trial are not too great. Such diversity (called also clinical heterogeneity) may distort the results to a significant degree. Clinical heterogeneity (being actually arbitrary, medical assessment of similarity between particular studies) should be distinguished from statistical heterogeneity, being an answer to the question

TABLE 4 Jadad scale[25]

Question	Answer Yes/No	Score	Comments of the assessor
Was the study described as randomized?			
Was the study described as double-blind?			
Was information concerning patient lost from the study and the follow-up period provided?			
Should 1 point be added for a correct and properly described method of randomization?			
Should 1 point be added for a correct and properly described method of blinding?			
Should 1 point be subtracted for an incorrect method of randomization?			
Should 1 point be subtracted for an incorrect method of blinding?			
Total			

whether differences between the results of particular studies may be due to chance.

An example of a meta-analysis, in which the problem of clinical heterogeneity remains controversial, is (in our opinion) the paper by Nissen and Wolski mentioned above.[13] The authors pooled the results of all identified trials, in which rosiglitazone was used, regardless of population characteristics, the comparators and methodology of the studies. The studies included in that meta-analysis concerned both patients with type 2 diabetes mellitus and those with prediabetes; placebo, insulin and oral medications (administered in monotherapy or combination) were used as comparators; the trials differed with respect to duration of treatment with rosiglitazone and methods of endpoint assessment.

What should an author of a meta-analysis do, having found that the studies are statistically heterogeneous? There is no simple and universal answer, but the problem must not be ignored. Discussing possible solutions exceeds the scope of this article. In short, potential factors of heterogeneity should be evaluated prior to analysis – this may be done using a single or multifactorial meta-regression analysis.[28]

A similar role may be fulfilled by subgroup analysis, i.e. assessment whether the results in one subgroup differ significantly from the results in another subgroup. In each situation a statistical form of meta-analysis assuming heterogeneity of the results between the studies, i.e. a random effect model, may be considered.

SUMMARY Over the last years significant progress with respect to quality and reliability of systematic reviews and meta-analyses has been observed. In many countries they became a basis of applications for financing health technologies from public means and in most cases published guidelines are based on systematic reviews of available data. Like all research tools, also these require professional and appropriate application. Of equal importance is adequate education of the main groups of recipients of those results – physicians and scientists – in order to enable them to evaluate the results received, which was the objective of the authors of this article.

REFERENCES

1 http://www.ncbi.nlm.nih.gov/sites/entrez/

2 Druss G. Growth and decentralization of the medical literature: implications for evidence-based medicine. J Med Libr Assoc. 2005; 93: 499-501.

3 Sackett DL, Rosenberg WMC. The need for evidence-based medicine. J R Soc Med. 1995; 88: 620-624.

4 Ioannidis JP, Haidich AB, Pappa M, et al. Contopoulos-Ioannidis DG, Lau J. Comparison of evidence of treatment effects in randomized and nonrandomized studies. JAMA. 2001; 286: 821-830.

5 Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ. 1996; 312: 1215-1218.

6 Guide to the Technology Appraisal Process (reference N0514). http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/guide_to_the_technology_appraisal_process_reference_n0514.jsp.

7 Light RJ, Pillemer DB. Summing Up: The Science of Reviewing Research. Cambridge (MA): Harvard University Press, 1984. In: Rationale for protocols. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.0 February 2008. http://www.cochrane-handbook.org/.

8 Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med. 1997; 126: 376-380.

9 Thacker, Stephen B. Meta-analysis: a quantitative approach to research integration. JAMA. 1988; 259: 1685-1689.

10 Shrier I, Boivin JF, Platt RW, et al. The interpretation of systematic reviews with meta-analyses: an objective or subjective process? BMC Med Inform Decis Mak. 2008; 21; 8: 19.

11 Hanefeld M, Cagatay M, Petrowitsch T, et al. Acarbose reduces the risk for myocardial infarction in type 2 diabetic patients: meta-analysis of seven long-term studies. Eur Heart J. 2004; 25: 10-16.

12 Van de Laar FA, Lucassen PLBJ, Akkermans RP, et al. Alpha-glucosidase inhibitors for type 2 diabetes mellitus. Cochrane Database of Systematic Reviews 2005, Issue 2. Art. No.: CD003639.

13 Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. N Engl J Med. 2007; 356: 2457-2471.

14 Richter B, Bandeira-Echtler E, Bergerhoff K, et al. Rosiglitazone for type 2 diabetes mellitus (Review). Cochrane Database Syst Rev. 2007; 3: CD006063.

15 Diamond GA, Bax L, Kaul S. Uncertain effects of rosiglitazone on the risk for myocardial infarction and cardiovascular death. Ann Intern Med. 2007; 147: 578-581.

16 Richardson WS, Wilson MC, Nishikawa J, et al. The well-built clinical question: a key to evidence-based decisions. ACP J Club. 1995; 123: A12-A13.

17 Counsell 1997. Counsell C. Formulating questions and locating primary studies for inclusion in systematic reviews. Ann Intern Med. 1997; 127: 380-387.

18 Defining types of outcomes: which outcome measures are most important? Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.0 February 2008. http://www.cochrane-handbook.org/

19 Gangji AS, Cukierman T, Gerstein HC, et al. A systematic review and meta-analysis of hypoglycemia and cardiovascular events: a comparison of glyburide with other secretagogues and with insulin. Diabetes Care. 2007; 30: 389-394.

20 Monami M, Lamanna C, Marchionni N, et al. Rosiglitazone and risk of cancer: a meta-analysis of randomized clinical trials. Diabetes Care. 2008; 31: 1455-1460.

21 Rao AD, Kuhadiya N, Reynolds K, et al. Is the combination of sulfonylureas and metformin associated with an increased risk of cardiovascular disease or all-cause mortality?: a meta-analysis of observational studies. Diabetes Care. 2008; 31: 1672-1678.

22 Willi C, Bodenmann P, Ghali WA, et al. Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. JAMA. 2007; 298: 2654-2656.

23 Health technology assessment guidelines. The Agency for Health Technology Assessment in Poland. Available from: www.aotm.gov.pl (September 2008).

24 Moher D, Cook DJ, Eastwood S et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet. 1999; 354: 1896-1900.

25 Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Controlled Clinical Trials. 1996; 17: 1-12.

26 Introduction to sources of bias in clinical trials. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.0 February 2008. http://www.cochrane-handbook.org/.

27 DCCT Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med. 1993. 329:977-986.

28 Meta-regression. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.0 February 2008. http://www.cochrane-handbook.org/.