New medicines in primary care

Dankers, Marloes

DOI:
10.33612/diss.809910284

2

Assessment of new medicines

2.1

Healthcare professionals' preferred efficacy endpoints and minimal clinically important differences in the assessment of new medicines for chronic obstructive pulmonary disease

Marloes Dankers
Marjorie H.J.M.G. Nelissen-Vrancken
Sara M.K. Surminski
Anke C. Lambooij
Tjard R. Schermer
Liset van Dijk

Front Pharmacol. 2020;10:1519.

ABSTRACT

Background

Registration authorities evaluate effects of new medicines for chronic obstructive pulmonary disease (COPD) on airway obstruction, dyspnea, health status and exacerbations. To establish clinical relevance, minimal clinically important differences (MCIDs) are used. The aim of this study was to investigate which efficacy endpoints and MCIDs healthcare professionals consider clinically relevant for new COPD medicines.

Materials and Methods

7,731 Healthcare professionals received an electronic questionnaire. Participants were asked for: 1) preferred efficacy endpoints for new COPD medicines and 2) cut-off values defining clinical relevance for forced expiratory volume in 1 sec (FEV_1), Transition Dyspnea Index (TDI) and St. George's Respiratory Questionnaire (SGRQ). Those cut-off values were compared to the MCIDs used by registration authorities, namely 100 ml for FEV_1, 1 unit for TDI and 4 units for SGRQ.

Results

227 Healthcare professionals responded to the questionnaire. Most preferred efficacy endpoints were exacerbations (51.0%), airway obstruction (46.9%) and health status (44.9%). Mean cut-off values for TDI and SGRQ were significantly higher than the corresponding MCIDs, mean differences 1.5 (95%CI = 1.3 – 1.8, $p > 0.001$) and 7.0 (95%CI = 5.1 – 8.8, $p < 0.001$), respectively. The mean cut-off value for FEV_1 was comparable to the MCID (mean difference 2.2, 95%CI = -19.9 – 24.3, $p = 0.84$).

Conclusions

Healthcare professionals largely agree with efficacy endpoints used for the evaluation of new COPD medicines. However, they seem to prefer higher cut-off values for clinical relevance for TDI and SGRQ than the registration authorities. Effects of new medicines on TDI and SGRQ that are considered clinically relevant by registration authorities do, therefore, not necessarily reflect healthcare professionals' perspectives on clinical relevance.

INTRODUCTION

New medicines are evaluated by registration authorities in order to assess their benefit-risk balance. The assessment of clinical relevance of new medicines by the registration authorities depends heavily on clinical trials. A common approach in clinical trials to investigate clinical effects is the use of patient-reported outcomes (PROs) [1]. PROs represent the patient perspective, quantifying the extent to which a disease impacts their health and functioning [2]. To establish the clinical relevance of a specific improvement on an endpoint, minimal clinically important differences (MCIDs) are used. An MCID is the smallest difference which patients perceive as beneficial and which would mandate a change in patient treatment [3]. Establishing whether an improvement on a clinical endpoint exceeds the MCID is a way to evaluate the clinical relevance of a (new) pharmacological treatment [4-6].

In the last decade, multiple new medicines for the treatment of chronic obstructive pulmonary disease (COPD) obtained market access [7-10]. Assessment of new COPD medicines includes the evaluation of effects on several efficacy endpoints, including airway obstruction, dyspnea, health status, and exacerbations. Frequently used parameters are forced expiratory volume in 1 sec (FEV_1), Transition Dyspnea Index (TDI) and St. George's Respiratory Questionnaire (SGRQ) for airway obstruction, dyspnea and health status, respectively [7-11]. Although FEV_1 is an objective endpoint and highly reproducible [12], it has a relatively poor correlation with symptoms [2,13]. PROs like dyspnea and health status might better reflect the impact of the disease on COPD patient's daily life. Improvements of 100 ml, one units and four units are validated MCIDs for FEV_1, TDI, and SGRQ, respectively [3,4,12,14]. Although these values are widely adopted, some debate about the acceptability of these MCIDs exists, especially for FEV_1 [3]. Values up to 140 ml are suggested as MCID for FEV_1 [12]. There is no clear MCID for the evaluation of exacerbations [15].

Since FEV_1, TDI, and SGRQ and their corresponding MCIDs are used by registration authorities for the evaluation of clinical efficacy of new medicines for the treatment of COPD, they are of crucial importance for the market access of these new medicines. Physicians and other healthcare professionals who prescribe (or advise about) those medicines have to rely on the assessment of new medicines by registration authorities. It is therefore of particular interest to know their opinions about the endpoints and MCIDs used in the assessment of clinical relevance of new medicines. Although expert-opinions can be included in the establishment of an MCID (in addition to the use of statistical and anchor-based approaches) [4,16], it is to our knowledge unknown how healthcare professionals assess the clinical importance of endpoints and their MCIDs

used for new COPD medicines. The aims of this study are therefore: 1) to investigate which efficacy endpoints healthcare professionals consider clinically relevant in the assessment of new medicines for COPD, and 2) to investigate which MCIDs healthcare professionals consider relevant for the frequently used endpoints FEV_1, TDI, and SGRQ.

MATERIALS AND METHODS

Background

Airway obstruction (measured by FEV_1), dyspnea (measured by TDI), and health status (measured by SGRQ) are important efficacy endpoints in the assessment of new medicines for COPD [11]. FEV_1 is the volume of air that is forcibly exhaled in the 1st second. The trough (pre-bronchodilator) FEV_1 is most commonly used in clinical trials evaluating the efficacy of COPD medicines [2]. A commonly used MCID for trough FEV_1 is 100 ml [3,4,12,14]. TDI is a validated evaluative instrument that measures changes in the severity of dyspnea by grading functional impairment, magnitude of task and magnitude of effort. Each parameter is graded from -3 to +3, adding up to a total score ranging from -9 to +9 [12]. The MCID is 1 unit [3,4,12,14]. SGRQ has been developed to measure health status in patients with respiratory disease. It is a self-administered questionnaire that measures health status in the subdomains symptoms, activity and impacts, with a total score ranging from 0 to 100 [2,12]. A difference of four units is considered clinically relevant [3,4,12,14].

Design

This investigation was part of a more extensive online survey about the opinion of Dutch healthcare professionals regarding the clinical relevance of new medicines for the treatment of type 2 diabetes mellitus (T2DM) and COPD. No ethical approval was needed. According to the Dutch legislation, neither obtaining informed consent nor approval by a medical ethics committee is obligatory for carrying out research among healthcare professionals that does not include patient data.

Participants

Participants for the online survey were obtained from the Customer relationship management (CRM) of the Dutch Institute for Rational Use of Medicine (IRUM). This CRM has multiple purposes, but is predominantly used for sending newsletters and information about the IRUMs activities. The CRM contained 7,731 email addresses of Dutch healthcare professionals (predominantly physicians, pharmacists, practice nurses, respiratory nurses, and diabetes nurses).

Data Collection

The invitation to fill out the questionnaire was sent by email (with a link to the questionnaire) on 15 November 2016. The online survey was closed two weeks later. All healthcare professionals received one reminder after 1 week. Participants did not receive a financial compensation, although every 10th participant was offered a free online accredited medicine course.

Questionnaire and Measurements

The full questionnaire consisted of 39 questions, among them 19 questions about new medicines for COPD. Responders were first asked for their profession. Next, they were asked whether they were involved in the management of patients with T2DM or COPD in their daily clinical practice. Only healthcare professionals working with COPD patients were asked to fill out the COPD section of the questionnaire.

The content of the COPD section of the questionnaire was based on the requirements for clinical trials for new COPD medicines, as described in the Guideline on clinical investigation of medicinal products in the treatment of COPD [11] and the MCIDs mentioned in the public assessment reports of new medicines for COPD [7-10].

The COPD section of the questionnaire consisted of three parts. The first part investigated the healthcare professionals preferred efficacy endpoints for the assessment of clinical relevance of new COPD medicines. Healthcare professionals were first informed about the need of efficacy endpoints in clinical trials and then asked which efficacy endpoints they considered clinically relevant for the assessment of new COPD medicines. All questions were open-ended in order to enhance the chance of getting reliable and sincere answers.

The second part investigated cut-off values for clinical relevance for FEV_1, TDI, and SGRQ. Healthcare professionals were informed about these endpoints and asked for a cut-off value for clinical relevance compared to placebo. All questions were open-ended. The MCIDs for FEV_1, TDI, and SGRQ were not mentioned before, in order to stimulate healthcare professionals to base their answer on their own clinical experience without being influenced by the validated MCIDs. Since TDI and SGRQ are not routinely used in daily clinical practice in the Netherlands, the questionnaire gave a brief description of these endpoints, including the range of those scales. This enabled responders to estimate a clinically relevant difference, regardless of their familiarity with those parameters. The third part about safety endpoints was not included in this analysis.

The online survey was programmed in NetQuestionnaire and pre-tested by three general practitioners, two pharmacists and three practice nurses for reasons of understandability and content. Since they experienced more difficulties with the COPD section compared to the T2DM section, the T2DM section was positioned before the COPD section to enhance the overall response rate.

Data Analysis

Responders were categorized by profession (physician, pharmacist, practice nurse, and other). The group "other" was excluded from further analysis. Responders who were both a physician and pharmacist were analysed as physicians. All different types of nurses (for example, practice nurses, respiratory nurses, and nurse practitioners) were analysed together as practice nurses.

After collection of the endpoints mentioned by the responders, six different categories of endpoints based on Global strategy for prevention, diagnosis and management of COPD 2019 report [17] were defined. Those categories were 1) exacerbations (including hospital admissions, infections, use of antibiotics and oral steroids), 2) airway obstruction (including parameters used to define airway obstruction like FEV_1), 3) health status (including quality of life, disease burden, Clinical COPD Questionnaire (CCQ) wellbeing and daily functioning), 4) respiratory symptoms (including dyspnea, cough, Medical Research Council (MRC) dyspnea scale, use of short-acting bronchodilators), 5) exercise intolerance (including condition, physical activity), 6) mortality, and 7) other (including morbidity, oxygen dependency or saturation, and adverse events). Subsequently, all answers were categorized by two researchers (based on consensus) and frequencies were calculated. Some healthcare professionals mentioned the Global Initiative for Chronic Obstructive Lung Disease (GOLD) status as preferred endpoint. Since GOLD includes airway obstruction, exacerbations and health status [17], each answer that mentioned GOLD was counted into these three categories. The categorization was verified by one independent researcher.

All cut-off values in open text fields were recoded into numeric variables. Impossible values for TDI and SGRQ (i.e., values exceeding the parameter range) were excluded. Based on expert opinions, FEV_1 values > 0 and < 1 ml and FEV_1 values > 1,499 ml were considered implausible and therefore also excluded, as were values in other units than asked (e.g., percentages instead of milliliters). Ranges (e.g., 50 – 75) were converted to averages (e.g., 67.5). Cut-off values for FEV_1, TDI and SGRQ were compared to the corresponding MCIDs by a one-sample T-test. Results were considered significant when $p < 0.05$.

All results were analysed with IBM SPSS Statistics 24.

RESULTS

A total of 556 responders (6.6%) started the questionnaire. Only healthcare professionals working with COPD patients in daily practice were included in this analysis. The final population consisted of 227 healthcare professionals (88 physicians, 107 pharmacists, and 27 practice nurses), resulting in a response rate of 2.9% (Figure 1). The group of 88 physicians contained two physicians who were also pharmacists and one dispensing general practitioner. Those were analysed solely as physicians. Among the group of 27 practice nurses were 21 practice nurses, one respiratory nurse, one physician assistant, one nurse practitioner, one trainee nurse practitioner, one geriatric nurse, and one practice nurse who was also a respiratory nurse and diabetes nurse. Those were analysed together as practice nurses.

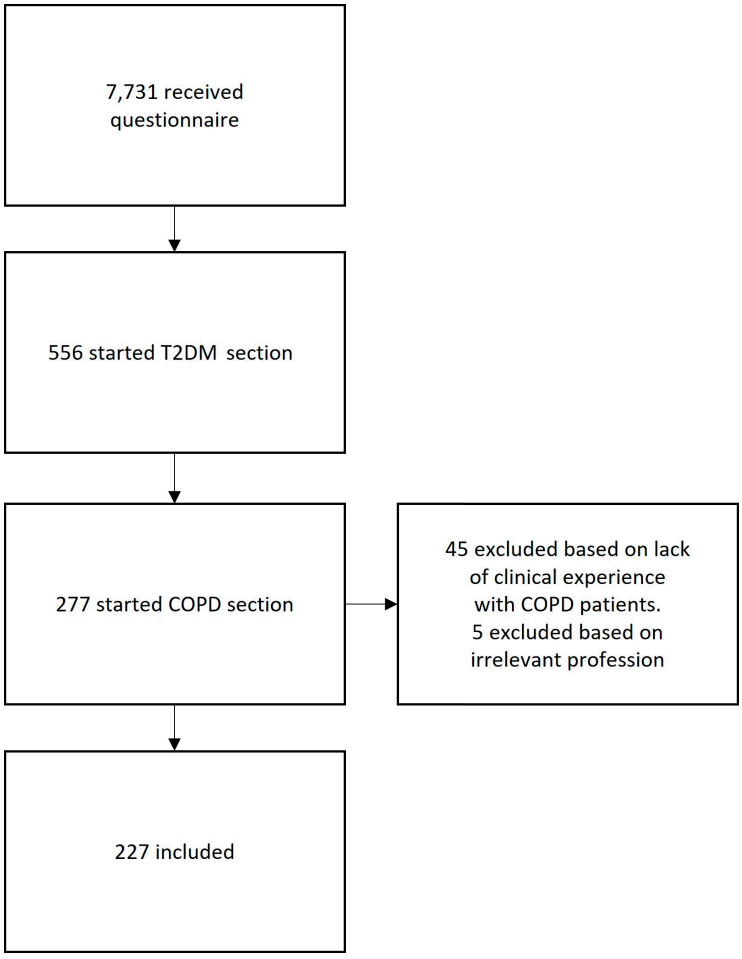

Figure 1. Response. COPD = chronic obstructive pulmonary disease, T2DM = type 2 diabetes mellitus.

Endpoints

196 healthcare professionals mentioned their efficacy endpoints of preference. The most frequently mentioned endpoints were exacerbations (51.0%), airway obstruction (46.9%), and health status (44.9%) (Table 1).

Table 1: Preferred efficacy endpoints in the evaluation of new COPD medicines.

Endpoints	Frequency
Exacerbations	51.0%
Airway obstruction[a]	46.9%
Health status	44.9%
Respiratory symptoms	30.6%
Mortality	23.0%
Exercise intolerance	9.2%
Other	11.2%

[a] *Two responders (1%) mentioned inspiratory capacity as most preferred endpoint. Since this is closely related to airway obstruction, those answers were counted into this category.*

MCIDs

Healthcare professionals were asked for a cut-off value that defined clinical relevance for FEV_1, TDI, and SGRQ. For both TDI and SGRQ, cut-off values according to healthcare professionals (2.5 and 11.0 units, respectively) were significantly higher than the MCIDs used by the European Medicines Agency (EMA) (one unit and four units, respectively) (Table 2). Mean differences for TDI and SGRQ compared to MCIDs were 1.5 (95% CI = 1.3 – 1.8, $p < 0.001$) and 7.0 (95% CI = 5.1 – 8.8, $p < 0.001$). Pharmacists mentioned the highest cut-off values for both TDI (2.8 units; followed by 2.4 units by physicians and 2.2 units by practice nurses) and SGRQ (12.2 units, followed by 11.5 units by practice nurses and 9.5 units by physicians).

The mean cut-off value for FEV_1 (102.2 ml) according to healthcare professionals was comparable to the MCID (100 ml), mean difference 2.2 ml (95% CI = -19.9 – 24.3, $p = 0.84$). Mean cut-off values according to pharmacists, physicians and practice nurses were 100.8 ml, 115.3 ml, and 72.1 ml, respectively. None of these cut-off values were significantly different from the MCID.

Table 2: Cut-off values for FEV_1, TDI, and SGRQ according to healthcare professionals, compared to the corresponding MCID.

		n	Mean (SD)	Mean difference compared to MCID (95%CI)	p
FEV_1 (ml) MCID = 100 ml	*All*	105	102.2 (114.1)	2.2 (-19.9 – 24.3)	0.84
	Physicians	44	115.3 (108.8)	15.3 (-17.8 – 48.3)	0.36
	Pharmacists	44	100.8 (133.6)	0.75 (-39.9 – 41.4)	0.97
	Practice nurses	17	72.1 (57.8)	-27.9 (-57.6 – 1.9)	0.064
TDI (unit) MCID = 1 unit	*All*	109	2.5 (1.4)	1.5 (1.3 – 1.8)	< 0.001*
	Physicians	44	2.4 (1.3)	1.4 (1.0 – 1.7)	< 0.001*
	Pharmacists	53	2.8 (1.5)	1.8 (1.4 – 2.2)	< 0.001*
	Practice nurses	12	2.2 (1.6)	1.2 (0.16 – 2.2)	= 0.027*
SGRQ (unit) MCID = 4 units	*All*	119	11.0 (10.1)	7.0 (5.1 – 8.8)	< 0.001*
	Physicians	50	9.5 (6.4)	5.5 (3.7 – 7.3)	< 0.001*
	Pharmacists	56	12.2 (11.9)	8.2 (5.0 – 11.3)	< 0.001*
	Practice nurses	13	11.5 (13.4)	7.5 (-0.56 – 15.6)	0.065

95%CI = 95% Confidence Interval, FEV_1 = forced expiratory volume in 1 sec, MCID = minimal clinically important difference, SD = standard deviation, SGRQ = St. George's Respiratory Questionnaire, TDI = Transition Dyspnea Index.
** Statistically significant different compared to MCID, based on one-sample T-test. Results were considered significant when $p < 0.05$.*

DISCUSSION

This study investigated which endpoints and MCIDs healthcare professionals considered clinically relevant for the evaluation of the efficacy of new medicines for COPD. Dutch healthcare professionals seem slightly more critical than registration authorities in the assessment of the clinical relevance of those new medicines. Although the preferred endpoints roughly correspond with the ones used in clinical trials, healthcare professionals prefer higher cut-off values for clinical relevance for TDI and SGRQ than the MCIDs used by registration authorities. This stricter view of clinical relevance is not seen for airway obstruction, since the average cut-off value for FEV_1 was comparable to the MCID. In general, physicians and practice nurses were less critical than pharmacists. This may display a difference in the clinical experience of healthcare professionals. Practice nurses and physicians more often see patients and measure endpoints like airway obstruction, dyspnea and health status than pharmacists. That might enhance their ability of estimating the expected medicine-induced improvement.

To our knowledge, this is the first study that specifically investigated the opinions of healthcare professionals about the endpoints and MCIDs used for the assessment of clinical efficacy of new COPD medicines. Expert opinions on the MCID for FEV_1 have,

however, been published before. The cut-off values for FEV_1 according to a small group of opinion leaders on this topic were generally higher than the MCID of 100 ml and thus also higher than the cut-off value for FEV_1 found in our investigation [3].

The differences between the cut-off values found in this investigation and established MCIDs might reflect the challenges with MCIDs as stated in other publications [3,14,18]. Factors such as heterogeneity in population and disease, trial duration, Hawthorne effects, withdrawal rates, and baseline disease severity may affect the size of benefit relative to the MCID [14,18]. It is therefore suggested that MCIDs should be used as an indicative value rather than as an absolute cut-off point [14]. The EMA, however, uses MCIDs to define the clinical relevance of new medicines. The cut-off values found in our study would have consequences for the evaluation of the clinical relevance of new COPD medicines. Multiple new (single-agent) inhalation medicines (aclidinium, glycopyrronium, indacaterol, and umeclidinium) for the treatment of COPD have been approved in Europe in the last decade. According to the EMA assessment of those medicines, roughly 50% of all improvements on FEV_1, TDI, and SGRQ exceeded the MCID and were thus considered clinically relevant [7-10]. When using the mean cut-off values found in the current study, instead of the MCIDs, none of the improvements on TDI and SGRQ would still have been clinically relevant. A new medicine that is considered "clinically relevant" by registration authorities does, therefore, not necessarily reflect healthcare professionals' views on clinical relevance. Since healthcare professionals have a stricter view of cut-off values for clinical relevance, defining clinical relevance by use of MCIDs might lead to overestimation of the expected treatment benefit.

Our results indicate that healthcare professionals consider exacerbations as the most important endpoint. Evaluation of the clinical importance of a reduction in exacerbations was not included in this investigation. Although evaluation of exacerbations is also part of the assessment procedure of new COPD medicines, there is no specific MCID used [7-10]. Defining an MCID for COPD exacerbations is problematic, because the impact of exacerbations is influenced dramatically by the used definition of (the severity of) an exacerbation and the influence of baseline status [15]. The use of exacerbation-free time instead of frequency (or severity) of exacerbations might better reflect the burden of exacerbations in COPD patients [19]. Future work should reveal the clinical relevance of a reduction in incidence or severity of exacerbations, or the increase of exacerbation-free time.

This study only included the assessment of the efficacy of new medicines for COPD. This is only a part of the assessment of new medicines, since safety and ease of use are also of clinical importance [7-10]. Our investigation did also not include patient

preferences on endpoints and MCIDs. A comprehensive literature review of patient preferences for the management of COPD revealed that the most important issues to patients with severe disease were symptom control, impact of disease on daily life, and positive relationship with the primary caregiver [20]. Another study reported the most reported ideal treatment factors based on interviews with 72 patients with asthma or COPD. These patients mentioned improved sleep, speed of action, and length of relief as most important aspects of treatment [21]. Patients perspectives on MCIDs are to our knowledge still unknown.

This investigation was meant as a first study to explore the opinion of different healthcare professionals (physicians, pharmacists, and practice nurses) on clinical relevance of endpoints and cut-off values. Since this study is based on the opinion of healthcare professionals working with COPD patients in daily practice, it provides a clear view of how clinical relevance of new medicines is considered in the daily practice of physicians, pharmacists and practice nurses. The main strength of this investigation is the exploratory and open character which was stimulated by the questionnaire with open-ended answers. There are, nonetheless, some limitations of this study. First, the results cannot be generalized to all Dutch healthcare professionals. Since the IRUM's CRM was used, only healthcare professionals who were somehow interested in pharmaceutical care were included in this study. Second, approximately half of the responders were pharmacists. In general, they will have less clinical experience than physicians and practice nurses. The overrepresentation of pharmacists could have influenced the mean cut-off values. However, the analysis of the cut-off values per profession showed that the results of the different professions were generally in line with each other. Third, the response rate was poor, with only 6.6 percent response to the general questionnaire and 2.9 percent to the COPD section. The number of responders that completed the questions about cut-off values was even lower. There are some possible explanations for this. Since the healthcare professionals were enrolled in a CRM instead of a research panel, they did not routinely participate in investigations and were not used to be approached for research via this panel. Another contributing factor to the low response was the fact that this research was part of a more extensive investigation towards the opinion of healthcare professionals about new medicines. A substantial number of healthcare professionals dropped out before the COPD section. However, unless the poor response rate, there was still a considerable number of healthcare professionals available for analysis. Fourth, the questions about cut-off values for TDI and SGRQ might have been fairly difficult to answer, since these instruments are not routinely used in Dutch clinical practice. This seems to be reflected by the wide range of cut-off values for clinical relevance. To maximize the probability of getting reliable results, the questionnaire referred to the range of scores for TDI and SGRQ. This enabled

healthcare professionals unfamiliar with these scales to estimate a clinically relevant difference. However, this does not completely rule out the possibility of inaccuracy in the mentioned cut-off values for TDI and SGRQ.

Despite these limitations, this study suggests that healthcare professionals are more critical than registration authorities in defining the clinical relevance of the efficacy of new medicines for the treatment of COPD. Larger and more representative ad hoc trials are needed to focus the topic and confirm these results. In the meantime, the established MCIDs should be used with caution, since new medicines that exceed the MCID do not necessarily meet the expectations of clinical relevance according to healthcare professionals. Defining clinical relevance by using MCIDs might, therefore, lead to overestimation of the expected treatment benefit of new COPD medicines by healthcare professionals.

REFERENCES

1. Hedayat A. S., Wang J., Xu T. (2015). Minimum clinically important difference in medical studies. Biometrics 71 (1), 33–41. 10.1111/biom.12251
2. Jones P., Miravitlles M., van der Molen T., Kulich K. (2012). Beyond FEV1in COPD: a review of patient-reported outcomes and their measurement. Int. J. Chron. Obstruct. Pulmon. Dis. 7, 697–709. 10.2147/COPD.S32675
3. Donohue J. F. (2005). Minimal clinically important differences in COPD lung function. COPD 2 (1), 111–124. 10.1081/COPD-200053377
4. Make B. (2007). How can we assess outcomes of clinical trials: The MCID Approach. COPD: J. Chronic. Obstruct. Pulmon. Dis. COPD. 4 (3), 191–194. 10.1080/15412550701471231
5. Koynova D., Lühmann R., Fischer R. (2013). A framework for managing the minimal clinically important difference in clinical trials. Ther. Innov. Regul. Sci. 47 (4), 447–454. 10.1177/2168479013487541
6. Bartels R. H. M. A., Donk R. D., Verhagen W. I. M., Hosman A. J. F., Verbeek A. L. M. (2017). Reporting the results of meta-analyses: a plea for incorporating clinical relevance referring to an example. Spine J. 17 (11), 1625–1632. 10.1016/j.spinee.2017.05.019
7. CHMP (2009). Assessment report: Onbrez Breezhaler, INN: indacaterol, Procedure No. EMEA/H/C/001114. (December). Available from: http://www.emea.europa.eu.
8. CHMP (2012). Assessment report: Eklira Genuair, INN: aclidinium bromide. Procedure No. EMEA/H/C/002211. (May). Available from: http://www.emea.europa.eu.
9. CHMP (2012). Assessment report: Seebri Brezhaler, INN: glycopyrronium bromide. Procedure No. EMEA/H/C/002430. (August). Available from: http://www.emea.europa.eu.
10. CHMP (2014). Assessment report: Incruse, INN: umeclidinium bromide. Procedure No. EMEA/H/C/002809/0000. (February). Available from: http://www.emea.europa.eu.
11. CHMP (2012). Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD). EMA/CHMP/700491/2012. 10.1183/09031936.00099306
12. Cazzola M., MacNee W., Martinez F. J., Rabe K. F., Franciosi L. G., Barnes P. J., et al. (2008). Outcomes for COPD pharmacological trials: from lung function to biomarkers. Eur. Respir. J. 31 (2), 416–469. 10.1007/164_2016_70
13. Singh D. (2017). Evaluation of new drugs for asthma and COPD: Endpoints, biomarkers and clinical trial design. Handb. Exp. Pharmacol. 237, 243–264. 10.3109/15412555.2012.733463
14. Jones P. W., Beeh K. M., Chapman K. R., Decramer M., Mahler D. A., Wedzicha J. A. (2014). Minimal clinically important differences in pharmacological trials. Am. J. Respir. Crit. Care Med. 189 (3), 250–255. 10.1097/00002281-200203000-00006
15. Chapman K. R., Bergeron C., Bhutani M., Bourbeau J., Grossman R. F., Hernandez P., et al. (2013). Do we know the minimal clinically important difference (MCID) for COPD exacerbations. COPD 10 (2), 243–249. 10.3109/15412555.2012.733463
16. Beaton D. E., Boers M., Wells G. A. (2002). Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. Curr. Opin. Rheumatol. 14 (2), 109–114. 10.1081/COPD-200050649
17. Global Initiative for Chronic Obstructive Lung Disease (GOLD) (2019). GOLD 2019 Global Strategy for the Diagnosis, Management and Prevention of COPD. 10.1038/s41533-018-0079-5
18. Kiley J. P., Ram J. S., Croxton T. L., Weinmann G. G. (2005). Challenges associated with estimating minimal clinically important differences in COPD - The NHLBI perspective. COPD 2 (1), 43–46. 10.2147/COPD.S82179

19. Boer L. M., Bischoff E. W., Borgijink X., Vercoulen J. H., Akkermans R. P., Kerstjens H. A. M., et al. (2018). Exacerbation-free time' to assess the impact of exacerbations in patients with chronic obstructive pulmonary disease (COPD): a prospective observational study. NPJ Prim. Care Respir. Med. 28 (1), 12. 10.1007/s12325-017-0557-0
20. Bereza B. G., Troelsgaard Nielsen A., Valgardsson S., Hemels M. E., Einarson T. R. (2015). Patient preferences in severe COPD and asthma: a comprehensive literature review. Int. J. Chron. Obstruct. Pulmon. Dis. 10, 739–744. 10.2147/COPD.S82179
21. Svedsater H., Roberts J., Patel C., Macey J., Hilton E., Bradshaw L. (2017). Life impact and treatment preferences of individuals with asthma and chronic obstructive pulmonary disease: results from qualitative interviews and focus groups. Adv. Ther. 34 (6), 1466–1481. 10.1007/s12325-017-0557-0

2.2

Alignment between outcomes and minimal clinically important differences in the Dutch type 2 diabetes mellitus guideline and healthcare professionals' preferences

Marloes Dankers
Marjorie H.J.M.G. Nelissen-Vrancken
Bertien H. Hart
Anke C. Lambooij
Liset van Dijk
Aukje K. Mantel-Teeuwisse

Pharmacol Res Perspect. 2021;9(3):e00750.

ABSTRACT

To evaluate the clinical benefit of new medicines for type 2 diabetes mellitus (T2DM), the Dutch guideline committee T2DM in primary care established the importance of outcomes and minimal clinically important differences (MCIDs). The present study used an online questionnaire to investigate healthcare professionals' opinions about the importance of outcomes and preferences for MCIDs. A total of 211 physicians, pharmacists, practice nurses, diabetes nurses, nurse practitioners and physician assistants evaluated the importance of mortality, macro- and microvascular morbidity, HbA1c, body weight, quality of life, (overall) hospital admissions and severe and other hypoglycemia on a 9-point scale. All outcomes were considered critical (mean scores 7-9), except for body weight and other hypoglycemia (mean scores 4-6). Only HbA1c and hospital admissions were valued differently by the guideline committee (not critical). Other relevant outcomes according to the respondents were adverse events, ease of use and costs. Median MCIDs were 4 mmol/mol for HbA1c (guideline: 5 mmol/mol) and 3 kg for body weight (guideline: 5 kg weight gain and 2,5 kg weight loss). Healthcare professionals preferred relative risk reductions of 20% for mortality (guideline: 10%) and macrovascular morbidity (guideline: 25%) and 50% for other hypoglycaemia (guideline: 25%). The MCID of 25% for microvascular morbidity, hospital admissions and severe hypoglycaemia corresponded to the guideline-MCID. Healthcare professionals' preferences were thus comparable to the views of the guideline committee. However, healthcare professionals had a stricter view on the importance of HbA1c and hospital admissions and the MCIDs for mortality and other hypoglycemia.

INTRODUCTION

The last two decades, new pharmacological treatments have become available for the treatment of type 2 diabetes mellitus (T2DM), including dipeptidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide-1 (GLP-1) receptor agonists and sodium-glucose-cotransporter 2 (SGLT2) inhibitors. Most of these drugs have found their ways into national and international clinical guidelines [1,2].

To evaluate new pharmacological treatments, guideline committees have to specify the criteria the medicines have to meet. Therefore, the importance of outcomes and cut-off values for a clinical benefit on these outcomes have to be established. The importance of outcomes can be scored according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach. GRADE recommends the use of a 9-point scale. A score of 1 – 3 indicates limited importance, 4 – 6 important, but not critical, and 7 – 9 critical importance [3]. Subsequently, cut-off values for a clinical benefit can be defined [4]. Those cut-off points, also known as minimal clinically important differences (MCIDs) or minimal important differences (MIDs), are used to evaluate the clinical relevance of a difference between two treatments [5].

For the evaluation of pharmacological treatments for T2DM, a considerable number of outcomes can be relevant, varying from direct outcomes for clinical efficacy (e.g. mortality) to surrogate outcomes (e.g. HbA1c), safety outcomes (e.g. hypoglycemia) and patient reported outcomes (PROs) like quality of life [6-10]. Validated MCIDs are not available for T2DM medicines. The decisions about importance of outcomes and MCIDs in treatment guidelines are therefore based on expert opinions and guideline committee consensus [4,11,12].

The abovementioned approach for the definition of importance of outcomes and corresponding MCIDs was also followed in het process of updating the Dutch clinical guideline for the treatment of T2DM in primary care in 2018 [4]. The outcomes mortality, macrovascular and microvascular morbidity, HbA1c, body weight, quality of life, (overall) hospital admissions, severe hypoglycemia, other hypoglycemia (not specified, mild or modest) and other adverse events were evaluated by the guideline committee for their relative importance. Subsequently, the guideline committee established the MCIDs for those outcomes [4]. The MCIDs were based on previously defined MCIDs in other national and international guidelines [2,11,12], non-specific thresholds for relative risks and standardized mean differences (SMD) provided by GRADE [13] and expert opinion in the guideline committee [4]. An overview of outcomes, their relative importance and MCIDs used during the Dutch T2DM guideline development can be found in Table 1.

Table 1: Overview of importance and MCIDs of outcomes used in the update of the T2DM guideline [2].

Outcomes	Importance[a]	Cut-off point for clinical relevance	MCID based on
All-cause mortality	Critical	RRR 10% (RR < 0,9 or RR > 1,1)	Expert opinion guideline committee
Macrovascular morbidity	Critical	RRR 25% (RR < 0,75 of RR > 1,25)	GRADE [13]
Microvascular morbidity	Critical	RRR 25% (RR < 0,75 or RR > 1,25)	GRADE [13]
Quality of life	Critical	Every statistically significant difference or SMD = 0,5	GRADE [13]
Severe hypoglycemia	Critical	RRR 25% (RR < 0,75 of RR > 1,25)	GRADE [13]
Other adverse events	Critical or important[b]	Every statistically significant difference or SMD = 0,5	GRADE [13]
Hospital admissions	Important	RRR 25% (RR < 0,75 of RR > 1,25)	GRADE [13]
Change in HbA1c	Important	0,5% or 5 mmol/mol	NICE guideline Type 2 diabetes in adults: management [11]
Change in body weight	Important	5% in case of both treatments cause weight gain 2,5% in case of one treatment causes weight loss and the other causes weight gain (or had a neutral effect on weight)	Dutch guideline T2DM in secondary care [12]
Other hypoglycemia (not specified or mild or modest)	Important	RRR 25% (RR < 0,75 of RR > 1,25)	GRADE [13]

MCID = minimal clinically important difference, RR = relative risk, RRR = relative risk reduction, SMD = standardized mean difference, T2DM = type 2 diabetes mellitus.

[a] Minor differences existed in relative importance between different healthcare questions. The importance shown is based on the importance for most healthcare questions.

[b] Depending on the severity of the adverse event.

The treatment recommendations in the final guideline heavily depend on the classification of importance of outcomes and MCIDs. Since the final guideline is leading for the treatment choices healthcare professionals in primary care make, it is of particular interest to know the degree of alignment between the guideline committee and the end users of the guideline. There is limited or no evidence concerning the views of healthcare professionals about the importance of outcomes and MCIDs used in guideline development or in the evaluation of blood glucose lowering drugs. Therefore, the aim of this study was to investigate healthcare professionals' opinions about the importance of outcomes and preferences for MCIDs used in the evaluation of new medicines for the T2DM guideline.

MATERIALS AND METHODS

Design

An online questionnaire was developed to investigate healthcare professionals' opinions about outcomes and MCIDs used in the evaluation of new T2DM medicines. According to the Dutch legislation, neither obtaining informed consent nor approval by a medical ethics committee is obligatory for conducting research among healthcare professionals that does not include patient data. Therefore, no ethical approval was needed.

Participants

Participants for the online questionnaire were approached using the mailing list for newsletters of the Dutch Institute for Rational Use of Medicine (IRUM). The mailing list contained 12,115 email addresses of stakeholders in pharmaceutical care, such as healthcare professionals and policy makers. Since there was no information about the profession of the subscribers, the questionnaire was sent to all subscribers. The respondents were asked for their profession in the questionnaire. Therefore, the selection of relevant professions could be made afterwards.

Data collection

The invitation to fill out the questionnaire was sent by email with a link to the online questionnaire on 17 February 2020. All subscribers received one reminder after 10 days (27 February 2020). The online questionnaire was closed on 13 March 2020. Participants did not receive a financial compensation, although every 10th participant was offered a free online accredited course about the treatment of T2DM, which is part of the IRUM continuous medical education program.

Questionnaire and measurements

The questionnaire is available in Appendix 1. The content of the questionnaire was based on the outcomes and cut-off points for clinical relevance used during the development of the T2DM guideline [4]. The questionnaire was developed by the researchers and fine-tuned during several sessions. Thereafter, the questionnaire was pre-tested by six healthcare professionals (a general practitioner, a public pharmacist, a hospital pharmacist, a practice nurse and two diabetes nurses). Based on their suggestions, an open-ended question that asked for other relevant outcomes was added. As expected, the test panel experienced the most difficulties with the questions about MCIDs, especially about relative risks. To simplify these questions, some minor linguistic changes were made. Also, an option 'I do not know/no opinion' was added to all questions about MCIDs.

The final questionnaire was programmed in Enalyzer. The questionnaire consisted of 24 questions. Respondents were first asked whether they were actively involved in the management of T2DM patients in their daily clinical practice. Only healthcare professionals working with T2DM patients were asked to complete the questionnaire. They were asked to score the importance of the outcomes used for the evaluation of new T2DM medicines on a 9-point scale, assuming they were a member of a guideline committee. Respondents could also (optionally) mention other relevant outcomes. The questionnaire then explained the situation where a new treatment was compared to a control treatment. Respondents were asked which difference they would define as MCID. Because of the expected difficulty of estimating relative risks, the questionnaire stated a fictional situation where an absolute number of patients in the control group of 1.000 patients experienced the outcome. Respondents were asked which (absolute) number of outcomes in the treatment group would demonstrate a clinical relevant difference. A fictional example was given for clarification purposes. All questions were open-ended, but only reasonable values (based on expert opinion) were permitted. The last part of the questionnaire was used for validation purposes. The questionnaire mentioned the used MCIDs for clinical relevance for HbA1c and mortality in the Dutch guideline, and asked the respondents whether they agreed with these values. These responses were triangulated with the corresponding open ended answers.

Data analysis

Respondents were categorized by profession. Other professions than physicians, pharmacists, practice nurses, diabetes nurses, nurse practitioners and physician assistants (physician associates) were not included in this analysis, because they were not considered as end–users of the guideline who have either prescription authority (physicians, diabetes nurses, nurse practitioners and physician assistants) or a direct influence on prescription behaviour (practice nurses and pharmacists). No distinction was made between healthcare professionals in primary and secondary care.

Mean scores for importance of the different outcomes (on a 9-point scale) were calculated. Differences in the scores for importance between outcomes were compared by paired samples T-test, and differences between professions with One-way ANOVA. Results were considered statistically significant when $p < 0.05$. The other outcomes mentioned in the open-ended questions were categorized by two researchers (based on consensus). One independent researcher verified the categorization.

For the analysis of cut-off points, one highly unlikely value for body weight decrease (a difference of 90 kg) was excluded from further analysis. Respondents who found every difference relevant were assumed to support the lowest difference possible (1).

The distribution of the cut-off points was plotted for all variables and medians were calculated. All results were analysed with IBM SPSS Statistics 24.

RESULTS

Characteristics of healthcare professionals

A total of 394 respondents started the questionnaire, of whom 329 were healthcare professionals working with T2DM patients. Other professions than physicians, pharmacists, diabetes nurses, practice nurses, nurse practitioners and physician assistants were excluded (n = 83, predominantly healthcare assistants and nurses other than practice nurses). Another 35 respondents dropped-out before the questions about relevance of outcomes. Therefore, the final population consisted of 211 healthcare professionals, including 44 physicians (predominantly general practitioners), 55 pharmacists (predominantly community pharmacists), 69 practice nurses, 27 diabetes nurses, 14 nurse practitioners and two physician assistants. Data of nurse practitioners and physician assistants were combined in the analysis, due to the low number of respondents and the similarity in profession.

Table 2: Characteristics of respondents.

	Physicians (n = 44)	**Pharmacists (n = 55)**	**Practice nurses (n = 69)**	**Diabetes nurses (n = 27)**	**Nurse practitioners/ physician assistants (n = 16)**
Female sex	20 (46%)	38 (69%)	68 (99%)	25 (93%)	11 (69%)
Age (years)					
20 – 39	8 (18%)	22 (40%)	7 (10%)	0 (0%)	2 (13%)
40 – 59	27 (61%)	26 (47%)	41 (59%)	21 (78%)	10 (63%)
≥ 60	9 (21%)	7 (13%)	21 (30%)	6 (22%)	4 (25%)
Working experience (years)					
< 5	6 (14%)	9 (16%)	5 (7%)	0 (0%)	7 (44%)
5 – 9	6 (14%)	7 (13%)	10 (15%)	1 (4%)	4 (25%)
10 – 14	4 (9%)	6 (11%)	26 (38%)	5 (19%)	2 (13%)
15 – 19	4 (9%)	4 (7%)	21 (30%)	14 (52%)	1 (6%)
≥ 20	24 (55%)	29 (53%)	7 (10%)	7 (26%)	2 (13%)
Number of patients contacts per week					
< 5	21 (48%)	9 (16%)	2 (3%)	3 (11%)	11 (69%)
5 – 10	17 (39%)	9 (16%)	17 (25%)	4 (15%)	4 (25%)
11 – 20	5 (11%)	11 (20%)	32 (46%)	11 (41%)	1 (6%)
≥ 20	1 (2%)	26 (47%)	18 (26%)	9 (33%)	0 (0%)

The distribution of sex, age, years of working experience and number of patient contacts per week is shown in Table 2. The location of the practices was well distributed among the Netherlands. The majority of the physicians and pharmacists was well-experienced (≥ 20 years of working experience).

Relevance of outcomes

Healthcare professionals valued severe hypoglycemia as the most important outcome measure (mean score 8.30), followed by mortality (8.14) and quality of life (8.13) (Table 3). All outcomes were considered less important than severe hypoglycemia, except for mortality (p = 0.074). However, small differences did not affect the importance according to the GRADE scaling: other hypoglycemia and body weight were seen as important outcomes (score between 4 and 6), all other outcomes were of critical importance (score between 7 and 9).

Table 3: Mean (SD) importance of outcomes measures, scored on a 9-point scale.

Outcome measure	Mean score importance (SD)	Importance according to GRADE scaling
Severe hypoglycemia	8.30 (0.818)	Critical
Mortality	8.14 (1.032)	Critical
Quality of life	8.13 (0.991)*	Critical
Macrovascular morbidity	8.00 (0.933)**	Critical
Microvascular morbidity	7.95 (0.911)**	Critical
Hospital admissions	7.65 (1.104)**	Critical
HbA1c	7.04 (1.388)**	Critical
Other hypoglycemia[a]	6.64 (1.625)**	Important
Body weight	6.46 (1.360)**	Important

SD = standard deviation.
[a] Mild, modest or not-specified.
** P = 0.01.*
*** P < 0.001 (all compared to severe hypoglycemia).*

There were some differences in the assessment of importance of outcomes between professions (Figure 1). Diabetes nurses gave the highest scores for many outcomes, meaning that they valued outcomes more important than other professions. Physicians and pharmacists most often gave the lowest scores. Statistically significant differences (p < 0.05) between professions were found for all outcomes, except for mortality (p = 0.716) and quality of life (p = 0.138).

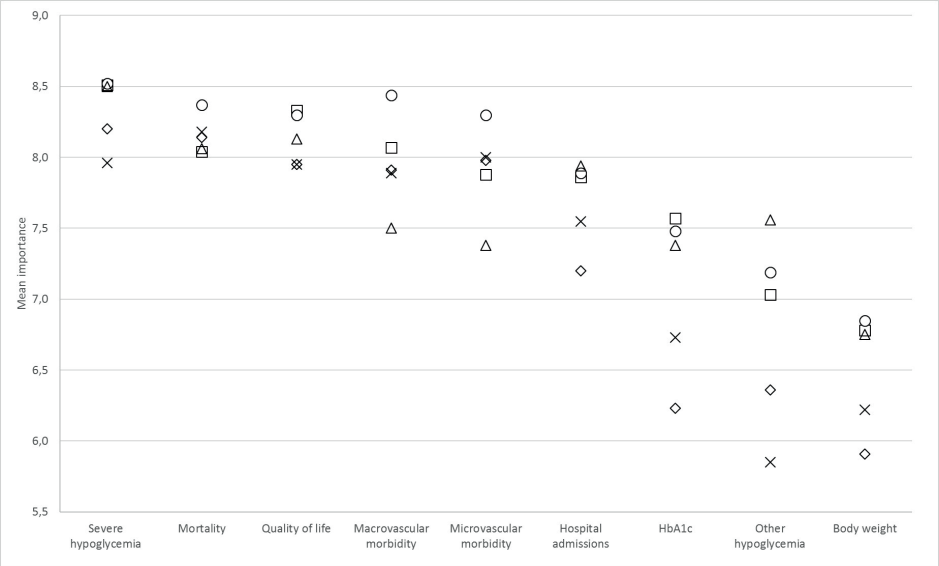

Figure 1: Mean importance of outcomes according to different professions. ◇ Physicians; X Pharmacists; □ Practice nurses; ○ Diabetes nurses; △ Nurse practitioners/physician assistants.

Of the 211 respondents, 114 healthcare professionals (54%) mentioned additional parameters they considered relevant in the assessment of blood glucose lowering drugs. Table 4 shows the outcome measures mentioned by at least two respondents. Adverse events (44.7%), ease of use (41.2%) and costs (10.5%) were most often mentioned.

Table 4: Other relevant outcomes mentioned by healthcare professionals.

Outcome measure	Number of respondents (%)
Adverse events[a]	51 (44.7)
Ease of use	47 (41.2)
Costs	12 (10.5)
Renal effects[b]	8 (7.0)
Effects on insulin use	4 (3.5)
Drug interactions	3 (2.6)
Glucose parameters other than HbA1c	2 (1.8)

[a] *Adverse events include some specific adverse events, like gastro-intestinal adverse events (n = 2), psychological adverse events (n = 2), lactate acidosis (n = 1) and fall risk (n = 1).*

[b] *Renal effects include renal adverse events as well as use by patients with renal impairment.*

MCIDs

MCIDs were investigated for HbA1c, body weight (increase as well as decrease), mortality, macrovascular and microvascular morbidity, hospital admissions, severe and other hypoglycemia. A considerable number of respondents found every difference clinically relevant or had no opinion (Table 5).

Table 5: Response on MCID-questions.

	N	Every difference relevant (%)	No opinion (%)
HbA1c	192	19%	21%
Body weight increase	191	29%	15%
Body weight decrease	185	25%	11%
Mortality	156	27%	19%
Macrovascular morbidity	156	22%	22%
Microvascular morbidity	156	21%	24%
Hospital admissions	156	24%	20%
Severe hypoglycemia	156	28%	16%
Other hypoglycemia	156	13%	27%

Respondents who had no opinion were excluded from further analysis. The results of the remaining healthcare professionals can be found in Figure 2. Median MCIDs according to healthcare professionals were 4 mmol/mol for HbA1c, 3 kg for weight increase as well as decrease, 20% for both mortality and macrovascular morbidity, 25% for microvascular morbidity, hospital admissions and severe hypoglycaemia and 50% for other hypoglycaemia.

For validity reasons, respondents were asked whether they agreed with the MCID used in the clinical guideline for HbA1c (5 mmol/mol) and mortality (RRR = 10%). Figure 3 shows the correspondence of this answer (x-axis) with the earlier preferred MCID, mentioned as open answer (y-axis). Although the answers roughly correspond, the wide range of answers (especially for mortality) shows that there was a considerable number of respondents whose answers were not in line. For example, a substantial amount of the respondents preferred an MICD < 10% for mortality according to the close-ended question, but previously mentioned an MCID > 10% in the open-ended question. Most likely, this indicates interpreting difficulties with the estimation of MCIDs, especially for RRRs.

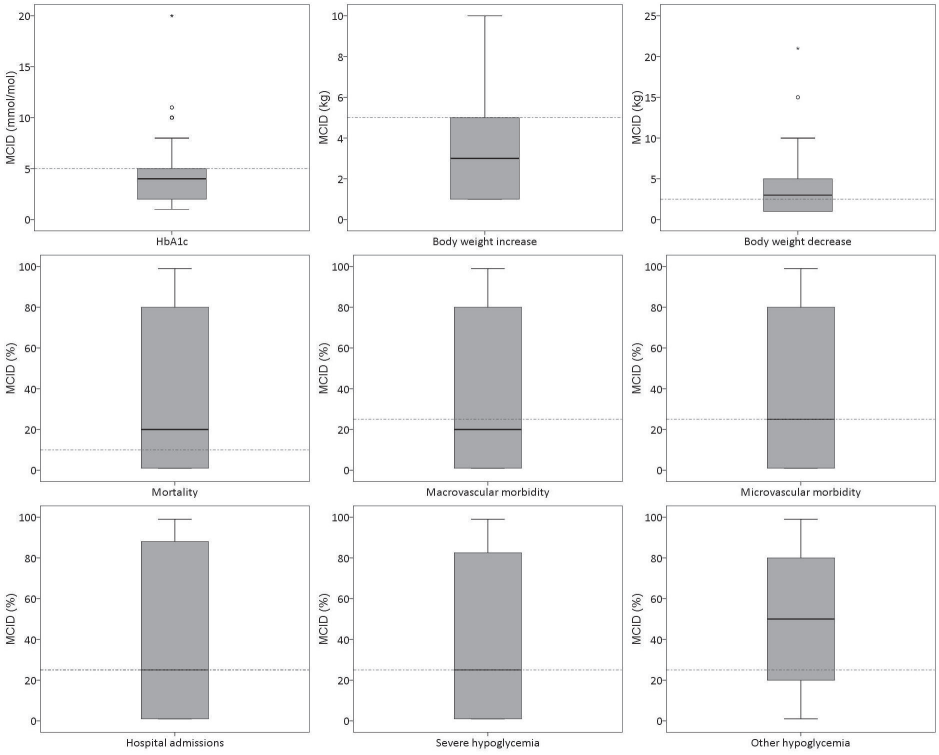

Figure 2 (A – I): Boxplots of MCIDs for outcomes. Dotted lines indicate the MCIDs used in guideline development.

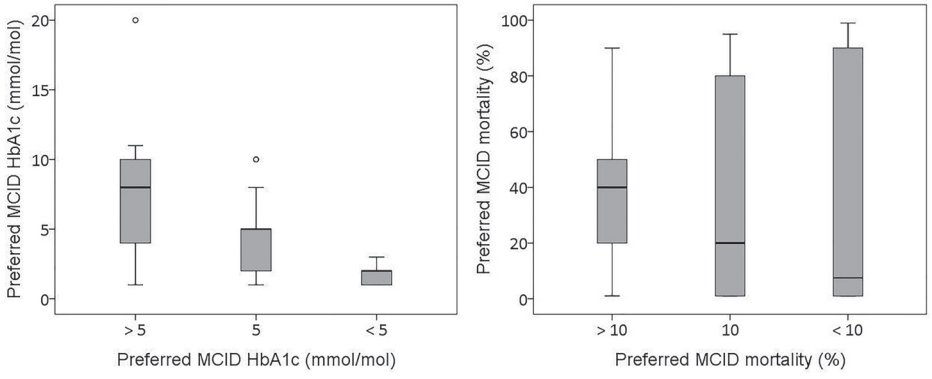

Figure 3 (A and B): Correspondence of close-ended (x-axis) and open-ended (y-axis) questions about preferred MCIDs for HbA1c and mortality.

DISCUSSION

In the evaluation of new T2DM medicines, healthcare professionals considered most outcomes used in the Dutch T2DM guideline in primary care as critically important. Exceptions were other hypoglycemia and change in body weight. Severe hypoglycemia was valued as the most important outcome, followed by mortality and quality of life. As additional parameters, adverse events, ease of use and costs were also seen as relevant. The preferred median MCIDs for HbA1c, body weight, macrovascular and microvascular morbidity, hospital admissions and severe hypoglycemia were comparable with the MCIDs used in the development of the Dutch T2DM guideline. For mortality and other hypoglycemia, healthcare professionals preferred higher median MCIDs [4]. However, this result should be interpreted with caution, because of the difficulties the respondents experienced with the estimation of MCIDs.

The views of healthcare professionals on importance of outcomes roughly correspond with the evaluation by the guideline committee. Compared to the guideline committee, only HbA1c and hospital admissions were valued differently (critical instead of important). The relevance of adverse events, ease of use and costs did also align. These outcomes were also considered during the process of the clinical guideline development, although at a later stage [4].

Remarkably, a safety outcome (severe hypoglycemia) was seen as most important, even more important than mortality and other efficacy parameters. Especially practice nurses, diabetes nurses and nurse practitioners/physician assistants valued the importance of severe hypoglycemia. Pharmacists gave the lowest scores for the importance of hypoglycemia (both severe and other). This difference might reflect the intensity of patient contacts among these professions. Healthcare professionals with many patient contacts will most likely have a more profound experience with hypoglycemia and thus are confronted with the impact of severe hypoglycemia on patients [14-16]. However, other explanations for the differences between professions cannot be excluded, since the distribution of sex, age and years of working experience were also markedly different between the professions.

The results of importance of outcomes are in line with a study by Mol et al. [17], that showed that physicians valued cardiovascular benefits of T2DM drugs as the most important aspect in making regulatory decisions. HbA1c, hypoglycemia and weight gain did also significantly affect physicians' choices [17]. A study by Gauthier et al. [16] however, showed that prescribers considered the overall efficacy in achieving glycemic control as the most important factor in choosing a blood glucose lowering drug if a patient

failed on metformin. Also, cost and insurance coverage, risk of hypoglycemia, weight gain, short- and long-term adverse events and quality of life were valued as important considerations. Clinical efficacy outcomes, like mortality and macro- and microvascular morbidity were barely mentioned [16]. The differences between the results of Gauthier et al. compared to our investigation and the study by Mol et al. [17] might be explained by the setting. Gauthier et al. investigated considerations in prescribing blood glucose lowering drugs to individual patients, while the investigation by Mol et al. and our study focused on decision-making at the population level in regulatory science and guideline development, respectively. In daily practice, decisions might be more influenced by short term outcomes on patient level, while clinical guidelines and regulatory agencies particularly focus on long term outcomes and population level [18,19]. Additionally, cultural differences and a shift towards valuing direct outcomes for clinical efficacy over surrogate outcomes during the last years could also have contributed [6,10]. Our study did not involve patients views on clinical relevance of T2DM drugs. However, their views have been investigated intensively elsewhere. Patients value glucose control, body weight, ease of use, hypoglycemia and other side effects important [14-16,20-22]. The views of patients—as well as the views of healthcare professionals—are mostly in line with those of regulators [17].

Our study also showed that healthcare professionals experience difficulties with estimating MCIDs, as was already concluded during the development of the questionnaire and the responses of the test panel. Despite the changes made for reasons of understandability, approximately 20 percent of the respondents had no opinion or did not answer the questions about MCIDs. Moreover the wide range of answers given, especially for RRRs, also indicate difficulties with the interpretation of these relative outcome measures [23,24]. However, the validation questions show that—despite the difficulties—there was reasonable alignment and the answers therefore give an indication about the estimation of MCIDs by healthcare professionals. The median MCIDs for HbA1c and body weight decrease were very close to the MCIDs used in guideline committees. The distinction made by the guideline committee between MCIDs for body weight decrease and increase was not seen in our results: the median MCID according to healthcare professionals was the same for both situations. The median MCIDs for other hypoglycemia was obviously higher (50%) than for mortality and macrovascular morbidity (20%) and microvascular morbidity, hospital admissions and severe hypoglycemia (25%). This also aligns with the establishment of relative importance of those outcomes, since other hypoglycemia was, among these outcome measures, also seen as the least important outcome. Due to these interpreting difficulties, no further analyses were performed on the MCIDs according to type of healthcare professional.

To the best of our knowledge, this is the first study that investigated the views of healthcare professionals about MCIDs used in the evaluation of T2DM medicines. However, we previously reported the correspondence of preferred outcomes and MCIDs for COPD medicines between healthcare professionals and regulatory agencies. Healthcare professionals preferred higher cut-off values for clinical relevance for COPD-related PROs than the MCIDs used by registration authorities [25]. In addition, the need for focus on clinical relevance in addition to statistical significance is often highlighted, in the conducting as well as reporting and interpretation of clinical trials [26,27]. The difficulties in the interpretation of risks and clinical relevance found in this study also highlights the need for education of healthcare professionals about the interpretation of clinical benefit of (new) medicines [23,24,28]. Moreover, the clinical relevance of new medicines can be over- or underestimated by healthcare professionals if the used outcomes and MCIDs in the evaluation of those medicines are not clearly communicated, especially since the views of healthcare professionals do not necessarily correspond with those of regulators and guideline committees.

This investigation was meant as a first study to explore the opinion of healthcare professionals on outcomes and MCIDs used in the evaluation of new medicines in the Dutch T2DM guideline in primary care. Since this study is based on the opinions of healthcare professionals working with T2DM patients, it provides a clear view of how clinical relevance of new medicines is considered in their daily practice. A main strength of this investigation is the exploratory and open character which was stimulated by the questionnaire with open-ended answers.

There are, nonetheless, some limitations of this study. First, the response rate seemed poor. This can be explained by the use of the mailing list for newsletters of the IRUM, which contains both email addresses of healthcare professionals and other stakeholders in pharmaceutical care. Since the profession of the subscribers was not known, it was not possible to target the invitation for the questionnaire. Although there was still a considerable number of 211 respondents, this approach might have limited the validity and generalizability of this study, also because only healthcare professionals that subscribed to the IRUMs newsletter and therefore will be interested in pharmaceutical care and IRUMs activities were included in this study. Second, no distinction could be made between healthcare professionals from primary and secondary care. Although most physicians and pharmacists were working in primary care, the work setting of the diabetes nurses and physician assistants/nurse practitioners was not known. Last, the questions about MCIDs, especially for RRRs were fairly difficult, as can be seen in the proportion of respondents that did not answer these questions and the wide range of answers. The examples given in the questionnaire for clarification purposes could

thereby have influenced the respondents. However, from the results of the validation questions it can be concluded that the majority of respondents interpreted the questions correctly, and the results for MCIDs can therefore be interpreted, albeit with caution.

This study must be seen as a first exploratory investigation towards the alignment of outcomes and MCIDs between the guideline committee T2DM and end users of the guideline. This study suggests that the views of healthcare professionals on the evaluation of importance of outcomes and MCIDs for the evaluation of new T2DM medicines are in line with the views in the guideline committee. However, HbA1c and hospital admissions were more important according to healthcare professionals and the MCIDs for mortality and other hypoglycemia were higher than the MCIDs used in the guideline. For those parameters, healthcare professionals were therefore more strict in defining clinical relevance than the guideline committee. Future research should confirm these results by the use of a larger representative group of healthcare professionals. In the meantime, clinical guideline committees should clearly communicate about how clinical relevance is established, so end users of the guideline can easily track the way new medicines were evaluated.

REFERENCES

1. Cosentino F, Grant PJ, Aboyans V, et al. 2019 ESC guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. Eur Heart J. 2020;41(2):255-323.
2. Nederland Huisartsen Genootschap (NHG) . NHG-Standaard diabetes mellitus type, vol. 2. [M01]. Utrecht: NHG; 2018.
3. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol. 2011;64(4):395-400.
4. Nederland Huisartsen Genootschap (NHG) . Totstandkoming en methoden NHG-Standaard Diabetes mellitus type 2 (M01). Utrecht: NHG; 2018.
5. Hedayat AS, Wang J, Xu T. Minimum clinically important difference in medical studies. Biometrics. 2015;71(1):33-41.
6. Rodriguez-Gutierrez R, McCoy RG. Measuring what matters in diabetes. JAMA. 2019;321(19):1865-1866.
7. Reaney M, Elash CA, Litcher-Kelly L. Patient reported outcomes (PROs) used in recent phase 3 trials for type 2 diabetes: A review of concepts assessed by these PROs and factors to consider when choosing a PRO for future trials. Diabetes Res Clin Pract. 2016;116:54-67.
8. Wieczorek A, Rys P, Skrzekowska-Baran I, Malecki M. The role of surrogate endpoints in the evaluation of efficacy and safety of therapeutic interventions in diabetes mellitus. Rev Diabet Stud. 2008;5(3):128-135.
9. Zannad F, Stough WG, Pocock SJ, et al. Diabetes clinical trials: Helped or hindered by the current shift in regulatory requirements? Eur Heart J. 2012;33(9):1049-1057.
10. Lipska KJ, Krumholz HM. Is hemoglobin A1c the right outcome for studies of diabetes? JAMA. 2017;317(10):1017-1018.
11. National Institute for Health and Care Excellence (NICE) . Type 2 diabetes in adults: management Clinical Guideline. [NG28]. London: NICE; 2015.
12. Nederlandse Internisten Vereniging (NIV) . Diabetes mellitus type 2 in de tweede lijn. Utrecht: NIV; 2018.
13. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. rating the quality of evidence–imprecision. J Clin Epidemiol. 2011;64(12):1283-1293.
14. Bøgelund M, Vilsbøll T, Faber J, Henriksen JE, Gjesing RP, Lammert M. Patient preferences for diabetes management among people with type 2 diabetes in denmark - a discrete choice experiment. Curr Med Res Opin. 2011;27(11):2175-2183.
15. Hauber AB, Mohamed AF, Johnson FR, Falvey H. Treatment preferences and medication adherence of people with type 2 diabetes using oral glucose-lowering agents. Diabet Med. 2009;26(4):416-424.
16. Gauthier B, Singh SR, Virani A, Staples H, Colbourne A. Perspectives and experiences of health care professionals and patients regarding treatments for type 2 diabetes. Can Pharm J (Ott). 2014;147(1):45-54.
17. Mol PG, Arnardottir AH, Straus SM, et al. Understanding drug preferences, different perspectives. Br J Clin Pharmacol. 2015;79(6):978-987.
18. Kaldjian LC. Patient care and population health: Goals, roles and costs. J Public Health Res. 2014;3(2):311.
19. Sox HC. Resolving the tension between population health and individual health care. JAMA. 2013;310(18):1933-1934.

20. Hayes RP, Bowman L, Monahan PO, Marrero DG, McHorney CA. Understanding diabetes medications from the perspective of patients with type 2 diabetes: Prerequisite to medication concordance. Diabetes Educ. 2006;32(3):404-414.
21. Mohamed AF, Zhang J, Johnson FR, et al. Avoidance of weight gain is important for oral type 2 diabetes treatments in sweden and germany: Patient preferences. Diabetes Metab. 2013;39(5):397-403.
22. Gelhorn HL, Stringer SM, Brooks A, et al. Preferences for medication attributes among patients with type 2 diabetes mellitus in the UK. Diabetes Obes Metab. 2013;15(9):802-809.
23. Brooks AP, Kibble SE, Alderson SA. The understanding of terms in evidence-based medicine: A pilot study. Eur Diabet Nurs. 2009;6(3):95-99.
24. Akobeng AK. Understanding measures of treatment effect in clinical trials. Arch Dis Child. 2005;90(1):54-56.
25. Dankers M, Nelissen-Vrancken MHJMG, Surminski SMK, Lambooij AC, Schermer TR, van Dijk L. Healthcare professionals' preferred efficacy endpoints and minimal clinically important differences in the assessment of new medicines for chronic obstructive pulmonary disease. Front Pharmacol. 2020;10:1519.
26. Sterne JA, Smith GD. Sifting the evidence-what's wrong with significance tests? Phys Ther. 2001;81(8):1464-1469.
27. Man-Son-Hing M, Laupacis A, O'Rourke K, et al. Determination of the clinical importance of study results. J Gen Intern Med. 2002;17(6):469-476.
28. Sedgwick P, Hall A. Teaching medical students and doctors how to communicate risk. BMJ. 2003;327(7417):694-695.

SUPPLEMENT

Questionnaire

Introduction

The *Instituut Verantwoord Medicijngebruik* (Institute for Rational Use of Medicine) studies the added value of new medicines in primary care. This questionnaire focuses on the criteria that guideline committees use to evaluate the clinical relevance (added value) of new medicines for type 2 diabetes mellitus (T2DM).

Answering the questionnaire takes approximately 10 minutes. Every 10th participant receives a free accredited online course on the treatment of T2DM. If you are interested in this course, please enter your e-mail address at the end of the questionnaire. We will only use your e-mail address to send the login code for the online course. All data will be processed anonymously.

Starting questions

We would first like to ask you some general questions about yourself and your working experience.

1. Are you involved in the daily treatment of patients with T2DM?
- o Yes
- o No (end of questionnaire)

2. What is your gender?
- o Woman
- o Man
- o Gender-neutral

3. What is your age?
- o Younger than 20 years
- o 20 to 39 years
- o 40 to 59 years
- o 60 years or older

4. What is your current profession?
- o Physician (forward to 5)
- o Pharmacist (forward to 6)
- o Practice nurse (forward to 7)
- o Diabetes nurse (forward to 7)
- o Nurse practitioner (forward to 7)
- o Physician Assistant (forward to 7)
- o Other, namely {open field} (forward to 7)

5. What is your specialization?
- o General practitioner with special interest in T2DM
- o General practitioner (including general practitioner trainee and dispensing general practitioner)
- o Internist
- o Other, namely {open field}

(forward to 7).

6. What is your specialization?
- o Community pharmacist (including community pharmacist specialist trainee)
- o Hospital pharmacist (including hospital pharmacist trainee)
- o Pharmacist in outpatient pharmacy in hospital
- o Other, namely {open field}

7. How many years of working experience in your current profession do you have?
- o Less than 5 years
- o 5 to 9 years
- o 10 to 14 years
- o 15 to 19 years
- o 20 years or more

8. What are the first 2 digits of the zip code for your working area? (We only use this answer to look at regional distribution).
{Open question: only answers between 10 and 99 allowed).

9. On average, how many T2DM patient contacts (for this condition) do you have per week?
o Less than 5
o 5 to 10
o 11 to 20
o More than 20

Outcome measures

The last years, new T2DM medicines have become available (DPP-4 inhibitors, GLP-1 receptor agonists and SGLT2 inhibitors). In order to develop a clinical guideline, the guideline committee first determines medicine relevant effects. We call this the outcome measures. For example, a new medicine for T2DM can be evaluated on the outcome measure 'mortality', but also on 'HbA1c' or 'hypoglycaemia'. The following questions concern your opinion on the importance of these outcome measures. In other words: should improvement of this outcome measure be included in the evaluation of a medicine?

You are a member of the guideline committee. According to you, how important are the effects of a blood glucose-lowering medicine on the following outcome measures? Please give your answer on a scale from 1 to 9, 1 meaning limited importance, 9 meaning critical importance.

- Mortality
- Macrovascular morbidity
- Microvascular morbidity
- HbA1c
- Body weight
- Quality of life
- Hospital admissions
- Severe hypoglycaemia
- Mild, moderate, or unspecified hypoglycaemia

{Answers on a scale of 1 - limited importance to 9 - critical}.

10. Are there other outcome measures you consider relevant when evaluating new medicine for T2DM? Which outcome measure(s)?
{Open question, not obligatory}.

Clinically relevant improvements

A guideline committee must also determine which difference in effect size between the new medicine and a control medicine is large enough to have added value for the patient. We call this difference or improvement clinically relevant. In the following questions you can indicate when you consider a difference to be clinically relevant.

11. HbA1c.

We compare a new medicine for T2DM with a control agent. The new medicine causes a larger HbA1c decrease in patients. What difference in HbA1c decrease between the control agent and the new agent do you consider clinically relevant?

For example, in the control group, the HbA1c decreases by 8 mmol/mol. In the group with the new medicine, the HbA1c decreases by 10 mmol/mol. Your answer is then 2 mmol/mol.

- Give your answer in mmol/mol (in whole numbers).
- Do you think every difference is clinically relevant? Then your answer should be "1".
- If you don't know or don't have an opinion, your answer should be '0'.

{Open question, only answers between 0 and 25 allowed}.

12. Body weight gain.

Both the new medicine and the control agents increase body weight. What difference in body weight gain between the control agent and the new agent do you consider clinically relevant?

For example, in the control group the body weight increases by 5 kg. In the group with the new agent, the body weight increases by 2 kg. Your answer is then 3 kg.

- Assume an average body weight of 100 kg.
- Assume that both the control agent and the new agent increase the body weight.
- Give your answer in kg (whole numbers).
- Do you think every difference is clinically relevant? Then your answer should be "1".
- If you don't know or don't have an opinion, your answer should be '0'.

{Open question, only answers between 0 and 99 allowed}.

13. Body weight decrease.

The new medicine reduces the body weight compared to the control agent. What difference in body weight reduction between the control and new agent do you consider clinically relevant?

For example, in the control group the body weight increases by 2 kg. In the group with the new medicine, the body weight decreases by 1 kg. Your answer is then 3 kg.

- Assume an average body weight of 100 kg.
- Assume the control agent increases or does not affect the body weight and the new agent decreases the body weight.
- Give your answer in kg (whole numbers).
- Do you think every difference is clinically relevant? Then your answer should be "1".
- If you don't know or don't have an opinion, your answer should be '0'.

{Open question, only answers between 0 and 99 allowed}.

Other outcomes

We compare a new T2DM medicine with a control agent.

1,000 patients use the control agent, of which 100 patients experience the outcome of interest.
1,000 other patients use the new medicine.

For how many outcomes in the group with the new medicine do you consider the difference to be clinically relevant?

For example, 100 out of 1,000 patients in the control group die. You think the difference in mortality is clinically relevant if only 20 out of 1,000 patients with the new medicine die. Your answer is then 20.

- Give your answer in number of outcomes (whole numbers).
- Do you think every difference is clinically relevant? Then your answer should be "99".
- If you don't know or don't have an opinion, your answer should be '0'.

14. Mortality {Open question, only answers between 0 and 99 allowed}.
15. Macrovascular morbidity {Open question, only answers between 0 and 99 allowed}.
16. Microvascular morbidity {Open question, only answers between 0 and 99 allowed}.
17. Hospital admissions {Open question, only answers between 0 and 99 allowed}.
18. Severe hypoglycaemia {Open question, only answers between 0 and 99 allowed}.

19. Mild, moderate or unspecified hypoglycaemia {Open question, only answers between 0 and 99 allowed}.

Cut-off values in guidelines

Guideline committees have established cut-off values for clinical relevance for some outcome measures. The following questions regard your opinion on these cut-off values for HbA1c and mortality.

20. For HbA1c, the Dutch guideline T2DM in primary care (2018) considers a difference of 5 mmol/mol clinically relevant. What do you think of this value?

- o Too low (I only consider a difference clinically relevant if it is greater than 5 mmol/mol)
- o Good
- o Too high (I consider differences less than 5 mmol/mol already clinically relevant)

21. For mortality, the Dutch guideline T2DM in primary care (2018) considers a relative risk reduction of 10% clinically relevant. What do you think of this value?

- o Too low (I only consider a difference clinically relevant if it is greater than 10%)
- o Good
- o Too high (I consider differences less than 10% already clinically relevant)

Final questions

You have reached the end of the questionnaire. Every 10th participant in the questionnaire receives a free accredited online course on the treatment of T2DM. If you are interested in this free course, please enter your e-mail address below.

22. Do you have any comments or questions regarding this questionnaire? {open question, not obligatory}.

23. What is your email address? {open question, not obligatory}.

Thank you for your cooperation! Click on ‘end survey’ to send your answers.