Statistical concerns about the study: “Risk factors for polypharmacy in older adults in a primary care setting: a cross-sectional study”

Dear editor

We read the article by Ersoy and Engin on the risk factors for polypharmacy in older adults in a primary care setting with great interest. We would like to add some comments that should improve the data interpretation in this large study.

Firstly, the authors noted that they assessed functionality by Activities of Daily Living and Instrumental Activities of Daily Living scales (ADL and IADL) with ADL consisting of five self-care measures, and IADL consisting of seven tasks. Scoring is undertaken using a 3-point ordinal scale, ranging from 0 to 2. The 0 point indicates inability, 1 indicates ability to do the task with aid, and 2 indicates ability to do it independently. The maximum score is 10 for the ADL and 14 for the IADL. The authors referred to the articles by Katz et al in 1963 and Lawton and Brody in 1969, respectively. However, Katz et al and Lawton and Brody’s assessments were not evaluated with five and seven items, respectively, and they did not use the 0–2 scale in the referenced articles. Instead, in the mentioned articles, ADL and IADL were assessed by six and eight items, respectively. Katz et al used an A to G scale to evaluate ADL and Lawton and Brody used a 0–1 scale to evaluate IADL. Accordingly, the maximum scores were not 10 and 14 but A (Katz et al for ADL) and 8 (Lawton and Brody for IADL), respectively. Furthermore, to our knowledge, the method the authors applied for evaluation of ADL and IADL has not been validated, yet. Thus, the methodology they used to assess ADL and IADL should be clarified and noted as limitation of the study.

Secondly, some statistical flaws were observed. The authors stated that they used Pearson correlation test to assess association between daily drug consumption (DDC) and continuous variables. However, the mean DDC was given as 4.63±3.51, with a very high SD value. This most probably suggests that the DDC parameter was a non-homogeneously distributed parameter. Hence, instead of Pearson correlation coefficient, Spearman Rho correlation should have been used. Similarly, while assessing the association between DDC and categorical variables such as presence of diabetes mellitus, metabolic syndrome, etc (as DDC seemed to be a non-homogenous parameter), the analyses should have been performed by Mann–Whitney U test instead of Student’s t-test.

Additionally, under the data analysis heading, they noted that linear regression analysis was performed to analyze the risk factors for DDC. However, the dependent parameters that should improve the data interpretation in this large study.

Secondly, some statistical flaws were observed. The authors stated that they used Pearson correlation test to assess association between daily drug consumption (DDC) and continuous variables. However, the mean DDC was given as 4.63±3.51, with a very high SD value. This most probably suggests that the DDC parameter was a non-homogeneously distributed parameter. Hence, instead of Pearson correlation coefficient, Spearman Rho correlation should have been used. Similarly, while assessing the association between DDC and categorical variables such as presence of diabetes mellitus, metabolic syndrome, etc (as DDC seemed to be a non-homogenous parameter), the analyses should have been performed by Mann–Whitney U test instead of Student’s t-test.

Additionally, under the data analysis heading, they noted that linear regression analysis was performed to analyze the risk factors for DDC. However, the dependent...
Aydin et al

variable should have been a homogeneously distributed variable for a given linear regression analysis. Again, DDC did not seem to be a homogenous variable as it had a very high SD value. Thus, statistical analyses should be re-performed to enable the readers to obtain the correct results with consequent data interpretation.

Disclosure
The authors report no conflicts of interest in this communication.

References
5. Hart A. Mann-Whitney test is not just a test of medians: differences in spread can be important. BMJ. 2001;323(7309):391–393.
Dear editor

We have read the letter by Aydin et al regarding our article “Statistical concerns about the study: risk factors for polypharmacy in older adults in a primary care setting: a cross-sectional study”.

The authors raised two main objections. The first one is about Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) scales.

Defining disability in functionally impaired people is complicated, especially in a heterogeneous population like older patients. Actually, there are plenty of functional ability instruments that have been developed to serve as a tool for objective assessment. In an effort to get more comprehensive and simpler tools, many ADL scales have been modified in terms of content and/or scoring system and Katz’s ADL is not an exception. While less than fifty ADL scales were in use in the last century, now we are talking about three-digit numbers. Our comprehensive geriatric assessment kit includes scales which were adopted from geriatric clinics of different medical schools in the early 2000s. It is known, for ADL measurements, that most authors used to select items from existing scales and make some adjustments for their clinical practice. This situation has raised concerns about validity and reliability. Under these circumstances, we failed to match our ADL and IADL scales with the references. That resulted in discordance between given references and our application. We sincerely apologize to readers and thank the reviewers for spotting our mistake. However, although this situation renders our ADL and IADL measurements debatable, it would not be justifiable to say the same thing of our results, as findings regarding ADL and IADL scales are only ancillary and contributory to the study. We respectfully affirm that our main results and conclusions are uninfluenced by these measurements.

As for the statistics, we again thank the authors for their question. This allows us to share our response to reviewers with the readers in general. We agree that in small samples with abnormal distribution, nonparametric tests must be used; even though there are studies contrary to this common rule. These studies reported that type 1 errors in two-group comparisons are well controlled in all sample sizes even if distribution is not normal. Nonetheless, our sample size was far from controversial limits.

Our data were not normally distributed but were roughly normal. This was one of our main concerns. Our statistical consultants have concluded that, in studies with such large sample sizes, parametric tests are robust even if the distribution is not normal, as type 1 error rates do not differ.

In sample sizes >100, parametric tests such as Student’s t-tests work well, even if the distribution is only approximately normal. Parametric tests have their advantages and we opted to use them.

Disclosure

The authors report no conflicts of interest in this communication.

References
